



# Reconnaissance de surfaces protéiques par des foldamères d'oligoamides aromatiques

Jeremie Buratto

## ► To cite this version:

Jeremie Buratto. Reconnaissance de surfaces protéiques par des foldamères d'oligoamides aromatiques. Cristallographie. Université de Bordeaux, 2014. Français. NNT : 2014BORD0003 . tel-01193295

**HAL Id: tel-01193295**

**<https://theses.hal.science/tel-01193295>**

Submitted on 4 Sep 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

PRÉSENTÉE A

**L'UNIVERSITÉ BORDEAUX 1**

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Par Jérémie BURATTO

POUR OBTENIR LE GRADE DE

**DOCTEUR**

SPÉCIALITÉ : Biochimie Structurale

## **RECONNAISSANCE DE SURFACES PROTEIQUES PAR DES FOLDAMERES D'OLIGOAMIDES AROMATIQUES**

Directeur de recherche : Dr. Bernard GALLOIS

Soutenue le : 7 Janvier 2014

Devant la commission d'examen formée de :

Dr Xavier MORELLI  
Dr Lionel MOUREY  
Dr Tom CESKA  
Dr Ivan HUC  
Dr Bernard GALLOIS

Directeur de recherche CNRS, Université Aix-Marseille  
Directeur de recherche CNRS, Université Toulouse 3  
Directeur de département, UCB Pharma  
Directeur de recherche CNRS, Université Bordeaux 1  
Directeur de recherche CNRS, Université Bordeaux 1

**Rapporteur**  
**Rapporteur**  
**Examineur**  
**Examineur**  
**Directeur de thèse**



# Résumé

Les interactions protéine - protéine sont au centre de nombreux processus biologiques, et représentent des cibles thérapeutiques pertinentes pour le traitement de certaines maladies. La conception de molécules antagonistes visant à inhiber ces interactions requiert la reconnaissance spécifique d'une des surfaces protéiques impliquées. Les foldamères de type oligoamides de quinoline constituent de bons candidats. Leur production et leur fonctionnalisation sont relativement aisées. Ils adoptent des structures hélicoïdales semblables à celles rencontrées dans les protéines. Grâce à différentes techniques biophysiques (CD, SPR, diffraction des rayons X), nous montrons que ces molécules sont aptes à reconnaître une surface protéique. Deux protéines cibles ont été choisies : l'interleukine 4 et l'anhydrase carbonique humaine II.

La stratégie retenue dans ce travail (ancrage du foldamère à la protéine via un bras espaceur) nous a permis d'obtenir des informations structurales sur les interactions foldamère – protéine avant toute optimisation de ces interactions. La première structure 3D d'un complexe foldamère de quinoline complexée à une protéine est décrite.

**Mots clés :** Reconnaissance de surface protéique, foldamères d'oligoamides de quinoline, interactions foldamère - protéine, interleukine 4 humaine, anhydrase carbonique humaine II.

# Abstract

Protein-protein interactions are a central issue in biological processes and represent relevant therapeutic targets for the treatment of diseases. The design of antagonistic molecules directed towards the disruption of these interactions requires the specific recognition of protein surfaces. Quinoline oligoamide foldamers present all the properties to reach that point. They are easily synthesized and fold into helices (similar to  $\alpha$  helices) which can be decorated. Thanks to biophysical studies (CD, SPR, RX diffraction), we demonstrate that these molecules are able to recognize protein surfaces. Two proteins have been studied: the human interleukin 4 and the human carbonic anhydrase II.

The applied strategy (keeping the foldamer close to the protein surface via a linker) allowed us to gather structural information about foldamer protein interactions before any strong binding is established. The first crystal structure between a protein and an aromatic amide foldamer is reported.

**Keywords:** Proteic surface recognition, quinoline oligoamide foldamer, foldamer - protein interactions, human interleukin 4, human carbonic anhydrase II.



---

# Remerciements

---

Ce travail de thèse a été réalisé au sein du laboratoire « Chimie et Biologie des Membranes et des Nano-objets » (CMBN, UMR 5248) dirigé par le Dr Erick Dufourc. Je lui exprime toute ma gratitude pour m'avoir accueilli au sein de son laboratoire.

C'est l'équipe « Structure et Activité des Macromolécules Biologiques » qui m'a très très bien accueilli et m'a immédiatement fait me sentir parmi les miens. Je tiens à remercier tout particulièrement et chaleureusement mon directeur de thèse (*The best one* !!) et directeur d'équipe, le Dr Bernard Gallois, pour la confiance qu'il m'a témoignée, pour ses conseils et son soutien sans faille dans la conduite de ce travail et dans la rédaction de ce manuscrit. Un grand MERCI !

Je souhaite sincèrement remercier le Dr Xavier Morelli et le Dr Lionel Mourey pour avoir accepté d'évaluer ce travail, ainsi que tous les autres membres du jury.

Un grand merci au Dr Ivan Huc, initiateur des travaux menés sur les foldamères d'oligoamides de quinolines et coordinateur actuel du projet FOLDAPPI, pour son soutien et ses conseils scientifiques toujours très avisés. Qu'il trouve ici l'expression de ma reconnaissance pour m'avoir considéré comme un membre de son équipe.

Je désire également remercier le Dr Tom Ceská pour son encadrement durant l'année passée au pays des *British* ainsi que tous les membres du labo d'UCB (Prash, Carl, Maja, Haiping, Andy et les autres).

Merci également à tous ceux (en particulier l'équipe de modélisation moléculaire dirigée par Michel Laguerre, ainsi que les massistes et les chimistes) qui ont participé à l'élaboration et au développement de ce projet multidisciplinaire passionnant qu'a été FOLDAPPI, et sans lesquels ce manuscrit n'aurait jamais vu le jour.

Un immense merci à l'infatigable Béatrice pour tous ses conseils, son soutien, les très nombreuses relectures et toutes les discussions scientifiques ou non que nous avons pu avoir. Merci également à Thierry pour son aide sans équivalent, ainsi qu'à Claude, Michel, Véronique et Jean. Je souhaite évidemment remercier Cinzia grâce à qui les foldamères ont pu si « facilement » être synthétisés, et pour les nombreux weekends passés à jongler entre SPR et CD. Merci également à Lucile, Laura, Simon, Christelle, Tiny et Marine pour leur travail en chimie, ainsi qu'à Brice pour son aide pendant les collectes aux synchrotrons.

Je ne dois pas oublier Guillaume, qui a partagé mon quotidien deux ans durant, qui a supporté mes « craquages » et mes coups de folie, et qui a toujours été présent. Bonne continuation à toi !! et surtout *smaklig måltid* !!!

Je remercierai enfin ma famille qui m'a toujours soutenu dans mes choix quels qu'ils soient depuis le début. Aussi je leur dédie ce travail.



---

# Sommaire

---

<b>Abréviations, unités et nomenclature</b>	13
<b>Avant-propos</b>	17
<b><u>Chapitre I</u> : Les interactions protéine-protéine et leur inhibition</b>	21
<b>Introduction</b>	23
<b>Partie 1. Interactions Protéine-Protéine (IPP)</b>	24
1.1. Définition et classification des IPP	24
1.2. Structure et description des IPP	26
1.2.1. Surfaces de contact	26
1.2.2. <i>Hot spots</i> ou points de contacts	27
1.3. Détection et analyse des IPP	29
<b>Partie 2. Inhibition des IPP</b>	31
2.1. Des objets biologiques pour inhiber les IPP	32
2.1.1. Oligonucléotides fonctionnels	32
2.1.2. Anticorps monoclonaux	33
2.1.3. Protéines miniatures	35
2.2. Des objets de synthèse pour inhiber les IPP	37
2.2.1. Systèmes à base de métaux	37
2.2.2. Mimes de feuillet $\beta$	38
2.2.3. Mimes d'hélices $\alpha$	39



<b>Partie 3. Les foldamères, des molécules qui se plient en quatre</b>	41
3.1. Différents squelettes ; différentes structures	42
3.1.1. Les foldamères aliphatiques	43
3.1.2. Les foldamères aromatiques	46
3.2. Les oligoamides de quinolines	49
<b>Conclusion</b>	52
 <b><u>Chapitre II : Inhibition des interactions protéine-protéine du complexe IL-4R<math>\alpha</math>/IL-4</u></b>	53
 <b>Introduction</b>	55
 <b>Partie 1. L'IL-4 : Fonctions, Structures et inhibition</b>	55
1.1. Physiologie de l'IL-4	55
1.1.1. Rôles et fonctions de l'IL-4	55
1.1.2. Mécanismes d'action de l'IL-4	56
1.2. Structures de l'IL-4 et du complexe IL-4R $\alpha$ /IL-4	58
1.2.1. Structure de l'IL-4	58
1.2.2. Le complexe IL-4R $\alpha$ /IL-4	59
1.2.3. Les complexes ternaires de type I et de type II	63
1.3. Stratégie d'inhibition du complexe	64
1.3.1. Par des anticorps	65
1.3.2. Par la mutation de protéines	65
1.3.3. Par des protéines miniatures	66
 <b>Partie 2. Nouvelle stratégie faisant appel aux foldamères</b>	67
2.1. Stratégie générale	67
2.2. Mutation, liaison chimique et espaceur	68

<b>Partie 3. L'adduit foldamère - IL-4 S16C</b>	69
3.1. Production de l'IL-4 mutée	69
3.1.1. Production et purification	69
3.1.2. Déglutathionylation	71
3.2. Ligation de foldamères sur l'IL-4 S16C	72
3.2.1. Le composé <b>A</b> , portant une fonction maléimide	72
3.2.2. Le composé <b>B</b> , portant une fonction thiol activée	74
3.2.3. Mauvais repliement de l'IL-4	78
3.2.4. Conclusion et discussion des résultats obtenus	79
3.3. Derniers résultats obtenus	80
 <b>Conclusion</b>	 81
 <b><u>Chapitre III</u> : L'anhydrase carbonique humaine : une preuve de concept</b>	 83
 <b>Introduction</b>	 85
 <b>Partie 1. L'anhydrase carbonique, une protéine modèle</b>	 85
1.1. Description et choix de la protéine	85
1.2. Stratégie appliquée pour la reconnaissance de la surface de la HCA-II par un foldamère	89
1.2.1. Stratégie générale	89
1.2.2. Conception du système d'ancrage	91
1.2.2.1. Position de l'espaceur sur l'inhibiteur	91
1.2.2.2. Structure du complexe <b>1</b> /HCA-II	92
1.2.2.3. Choix de l'espaceur et de la séquence du foldamère cible	94
1.2.2.4. Effet de l'addition de l'espaceur sur l'inhibiteur modifié	96
	9

<b>Partie 2. Les foldamères de première génération</b>	98
2.1. Bibliothèque	98
2.2. Evaluation des interactions foldamère/protéine	98
2.2.1. Principe du Dichroïsme Circulaire Induit (ICD)	98
2.2.2. Induction de chiralité du trimère <b>3</b> par la HCA-II	99
2.2.3. Induction de chiralité du pentamère <b>4</b> par la HCA-II	100
2.2.4. Induction de chiralité de l'heptamère <b>5</b> par la HCA-II	101
2.3. Tentatives de cristallisation	102
2.3.1. Cristallisation du complexe <b>3</b> /HCA-II	102
2.3.2. Cristallisation du complexe <b>4</b> /HCA-II	103
2.3.3. Cristallisation des heptamères <b>5</b> et <b>6</b> avec la HCA-II	103
<b>Partie 3. Les foldamères de seconde génération</b>	107
3.1. Bibliothèque	107
3.2. Identification d'un candidat potentiel par analyses biophysiques	108
3.2.1. Dichroïsme Circulaire Induit	108
3.2.2. Analyses par Résonance Plasmonique de Surface (SPR)	110
3.2.3. Evaluation du rôle joué par les chaînes latérales	111
3.2.3.1. Principe de la méthode utilisée	111
3.2.3.2. Impact des mutations réalisées	112
3.3. Structure du complexe <b>8</b> /HCA-II	113
3.3.1. Cristallisation du complexe <b>8</b> /HCA-II	113
3.3.2. Structure du complexe <b>8</b> /HCA-II	114
3.3.2.1. Collecte et analyse des données	114
3.3.2.2. Description générale, dimérisation du complexe	115
3.3.2.3. Contacts protéine/protéine	116
3.3.2.4. Les inhibiteurs interagissent avec leur site actif	118
3.3.2.5. Les foldamères interagissent entre eux...	120
3.3.2.6. ... et avec les protéines	121
<b>Conclusion</b>	123

## **Conclusion Générale et Perspectives** 125

## **Chapitre IV : Matériel et méthodes** 129

### **Partie 1. Biologie moléculaire et cellulaire** 131

1.1. Souches bactériennes utilisées 131

1.2. Plasmides et séquences utilisés 131

1.3. Quantification des fragments d'ADN 131

1.4. Mutagenèse dirigée 132

1.5. Transformation bactérienne 132

1.6. Milieux de culture 132

### **Partie 2. Biochimie** 133

2.1. Dosage des protéines 133

2.2. Electrophorèse des protéines par SDS-PAGE 133

2.3. Production et purification du mutant S16C de l'IL-4 134

2.4. Réduction enzymatique de l'adduit IL-4/Glutathion 135

2.5. Quantification des cystéines libres par le réactif d'Ellman 136

2.6. Production et purification de la HCA-II 137

### **Partie 3. Biophysique** 137

3.1. Dichroïsme circulaire induit (ICD) 137

3.2. Résonance plasmonique de surface (SPR) 138

### **Partie 4. Biologie structurale** 141

4.1. Modélisation moléculaire 141

4.2. Diffraction des rayons X 142

4.2.1. Cristallogenèse 142

4.2.1.1. Principe 142

---

4.2.1.2. Méthodes et techniques	144
4.2.1.3. Recherche des conditions et optimisation	145
4.2.1.4. Congélation des cristaux	147
4.2.2. Cristallographie	147
4.2.2.1. Diffraction et traitement des données	147
4.2.2.2. Résolution du problème des phases	150
4.2.2.2.1. Remplacement isomorphe	150
4.2.2.2.2. Diffusion anormale	150
4.2.2.2.3. Remplacement moléculaire	151
4.2.2.2.4. Amélioration des cartes de densité	152
4.2.2.3. Affinement et validation du modèle	152
 <b>Références Bibliographiques</b>	 155
 <b>Production Scientifique</b>	 165

# Abréviations, unités et nomenclature

Les abréviations et notations particulières utilisées dans ce manuscrit sont explicitées ci-dessous :

<b>ADN</b>	Acide DésoxyriboNucléique
<b>ADNc</b>	Acide DésoxyriboNucléique complémentaire
<b>APS</b>	Ammonium PerSulfate
<b>ARN</b>	Acide RiboNucléique
<b>CD</b>	<i>Circular Dichroism</i> (Dichroïsme Circulaire)
<b>C<math>\alpha</math></b>	Carbone alpha (d'un acide aminé)
<b>DLS</b>	<i>Dynamic Light Scattering</i> (Diffusion Dynamique de la Lumière)
<b>DMSO</b>	Diméthylsulfoxyde
<b>DO</b>	Densité Optique
<b>DTNB</b>	acide DiThiodiNitroBenzoïque
<b>DTT</b>	DiThioThrétol
<b>EDC</b>	1-Ethyl-3-(3-diméthylaminopropyl)carbodiimide
<b>EDTA</b>	acide EthyleneDiamineTétraAcétique
<b>EtOH</b>	Ethanol
<b>GSH</b>	Glutathion sous sa forme réduite
<b>GSSG</b>	Glutathion sous sa forme oxydée
<b>GuCl</b>	Chlorure de Guanidinium
<b>HCA-II</b>	<i>Human Carbonic Anhydrase II</i> (Anhydrase Carbonique Humaine II)
<b>HEPES</b>	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
<b>iBu</b>	<i>iso</i> Butyle
<b>IC<sub>50</sub></b>	Concentration d'un composé inhibant 50 % de l'effet observé
<b>ICD</b>	<i>Induced Circular Dichroism</i> (Dichroïsme Circulaire Induit)
<b>IL-4</b>	InterLeukine 4
<b>IPTG</b>	<i>iso</i> Propyl $\beta$ -D-1-thiogalactopyranoside
<b>ITC</b>	<i>Isothermal Titration Calorimetry</i> (Titration Calorimétrique Isotherme)
<b>K<sub>D</sub></b>	Constante d'affinité
<b>K<sub>i</sub></b>	Constante d'inhibition
<b>k<sub>off</sub></b>	Constante de dissociation
<b>k<sub>on</sub></b>	Constante d'association
<b>MAD</b>	<i>Multi-wavelength Anomalous Dispersion</i>
<b>MIR</b>	<i>Multiple Isomorphous Replacement</i>
<b>MIRAS</b>	<i>Multiple Isomorphous Replacement with Anomalous Scattering</i>
<b>MS</b>	<i>Mass Spectrometry</i> (Spectrométrie de Masse)
<b>MW</b>	<i>Molecular Weight</i> (Masse Moléculaire)
<b>NADPH</b>	Nicotinamide Adénine Dinucléotide Phosphate
<b>NHS</b>	N-Hydroxysuccinimide
<b>PBS</b>	<i>Phosphate Buffer Saline</i> (Tampon Phosphate Salin)

<b>PDB</b>	<i>Protein Data Bank</i>
<b>PEG</b>	Polyéthylène glycol
<b>pH</b>	Potentiel Hydrogène
<b>pHMB</b>	<i>para</i> -HydroxyMercuriBenzoate de sodium
<b>pI</b>	Point Isoélectrique
<b>rpm</b>	Rotation Par Minute
<b>RX</b>	Rayons X
<b>SDS</b>	Sodium Dodécyl Sulfate
<b>SIR</b>	<i>Single Isomorphous Replacement</i>
<b>SIRAS</b>	<i>Single Isomorphous Replacement with Anomalous Scattering</i>
<b>SPR</b>	<i>Surface Plasmon Resonance</i> (Résonnance Plasmonique de Surface)
<b>STE</b>	Saccharose Tris EDTA
<b>TEMED</b>	Trétraméthyléthylènediamine
<b>TGS</b>	Tris Glycine SDS
<b>TNB</b>	acide ThioNitroBenzoïque
<b>Tris</b>	2-amino-2-hydroxyméthyl-1,3-propanediol
<b>UV</b>	Ultraviolet
<b><math>\alpha, \beta, \gamma</math></b>	Angles formés par les trois axes de la maille élémentaire

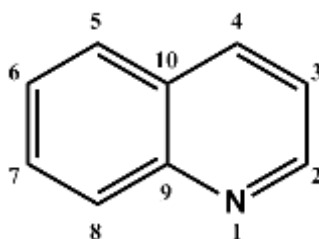
Les acides aminés au sein des séquences peptidiques sont notés par leur code à une ou trois lettres dans le texte. Ces codes sont rappelés ci-dessous :

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Méthionine	Met	M
Aspartate	Asp	D	Phénylalanine	Phe	F
Cystéine	Cys	C	Proline	Pro	P
Glutamate	Glu	E	Sérine	Ser	S
Glutamine	Gln	Q	Thréonine	Thr	T
Glycine	Gly	G	Tryptophane	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Les unités utilisées sont listées ci-dessous, avec leurs sous-multiples :

°C	Degré Celcius		
Å	Angström		
cal	Calorie		
Da	Dalton	c	centi ( $10^{-2}$ )
eq	Equivalent	m	milli ( $10^{-3}$ )
g	Gramme	μ	micro ( $10^{-6}$ )
h	Heure	n	nano ( $10^{-9}$ )
j	Jour	p	pico ( $10^{-12}$ )
L	Litre		
M	Molaire		
min	Minute		
mol	Mole		

La nomenclature concernant le monomère de quinoline (Q) est la suivante :







## Avant-propos

Il est communément admis que la vie a fait son apparition sur Terre il y a environ 4 milliards d'années. Depuis la formation des premières unités vivantes ou protocellules, l'évolution a sélectionné des êtres vivants de plus en plus complexes et perfectionnés, adaptés ou s'adaptant plus rapidement aux conditions environnementales. La cellule, unité constitutive et fonctionnelle du vivant, est composée (i) d'une ou plusieurs membranes délimitant les milieux intérieurs du milieu extérieur, (ii) de matériel génétique (ADN et/ou ARN), (iii) et de protéines, véritables machines moléculaires. Ainsi au fil des générations les protéines possédant la meilleure affinité, spécificité, constante catalytique furent sélectionnées. Les protéines sont au centre de toutes les fonctions cellulaires, elles possèdent un rôle dans la catalyse chimique, la régulation homéostatique, le transport, la composition structurale ou encore la motilité. Leurs actions et la régulation de leurs actions sont dépendantes des contacts qu'elles établissent avec d'autres protéines. Ainsi il serait plus juste de dire que ce sont les interactions protéine-protéine (IPP) qui sont au cœur de toutes les fonctions cellulaires. On parle d'interactome (Vidal *et al.* 2011, Veselovsky *et al.* 2002). Par exemple, la réponse immunitaire chez les mammifères repose sur la reconnaissance d'antigènes par des anticorps. De même l'interaction entre la tête de la myosine et une unité du filament d'actine permet la contraction musculaire. La transduction du signal de la surface de la cellule à son noyau est également contrôlée par des associations protéine-protéine, impliquant plus particulièrement les protéines G, en début de cascade.

Le génome humain contient environ 30 000 gènes et il est admis que plus de 10 000 protéines différentes sont présentes dans une cellule à un instant donné. Il a été estimé que plus de 80 % de ces protéines n'agissent pas seules mais au sein de complexes (Berggard *et al.* 2007). En raison de leur rôle central dans tous les mécanismes cellulaires et de leur nombre, les protéines et leurs interactions sont des cibles particulièrement intéressantes dans la conception de traitements thérapeutiques. La compréhension de ces interactions (Stites 1997) et leurs prédictions pourront enrichir les connaissances actuelles des mécanismes cellulaires qui régissent aussi bien les fonctions physiologiques que les maladies dues aux dérèglements de ces interactions, et ainsi promouvoir et accélérer la découverte de nouveaux agents thérapeutiques.

Ce manuscrit s'inscrit dans la démarche d'inhiber des IPP par la conception rationnelle d'objets synthétiques, adoptant des structures repliées remarquablement stables et prévisibles, les foldamères. Ce travail est l'aboutissement de trois années de recherche réalisées au sein de l'unité mixte de recherche du CNRS 5248 CBMN (Chimie et Biologie des Membranes et des Nano-objets) dans l'équipe du Dr Bernard Gallois, et au sein du laboratoire de recherche d'UCB Pharma à Slough en Angleterre. Il s'articule autour de 3 chapitres :

- Le premier explicite les caractéristiques qui font des IPP des cibles thérapeutiques porteuses d'espoir. Nous développerons également différentes méthodes d'inhibition des IPP utilisées à ce jour. Nous aborderons finalement plus en détails l'approche faisant appel aux foldamères.
- Le second chapitre est dédié à l'utilisation des foldamères de quinoline comme outils pour établir une reconnaissance spécifique de l'interleukine 4 (IL-4). Cette reconnaissance aura pour but d'inhiber l'interaction entre l'IL-4 et son récepteur. Cette application fait l'objet d'un projet européen, FOLDAPPI, impliquant des partenaires académiques et industriels, et qui m'a permis de poursuivre mes travaux au sein du laboratoire de recherche d'UCB Parma à Slough en Angleterre pendant une année.
- Dans le troisième chapitre, nous détaillons la stratégie utilisée pour concevoir un foldamère interagissant avec une protéine modèle, l'anhydrase carbonique humaine de type II (HCA-II).





## Chapitre I

Les interactions protéine-protéine et leur inhibition



## Introduction

Tous les grands processus biologiques impliquent des enzymes (siège de la catalyse de réactions) et des interactions protéine-protéine (IPP) en grand nombre. Ainsi le moyen le plus logique pour contrôler, améliorer, ou inhiber une fonction cellulaire serait d'agir sur une ou concomitamment plusieurs étapes mises en jeu dans cette fonction.

La communauté scientifique, et plus particulièrement les grands laboratoires pharmaceutiques, s'intéressent de plus en plus aux IPP. En effet, la très grande majorité des médicaments mis sur le marché sont des petites molécules qui agissent généralement comme des inhibiteurs d'enzyme, prenant la place du substrat, et empêchant/ralentissant la réalisation de la catalyse. Mais selon une étude récente (Verdine et Walensky 2007) basée sur l'analyse des structures de protéines déposées à la PDB, seules 3 000 des 30 000 protéines codées par le génome humain possèderaient un site actif compatible avec les petites molécules organiques synthétisées à des fins d'inhibition. Aussi, beaucoup d'efforts sont fait pour découvrir de nouvelles classes de molécules permettant d'inhiber les IPP. Les stratégies utilisant des anticorps (Liang *et al.* 2013, Chen *et al.* 2013) ou des peptides (Matthews *et al.* 2004) sont prometteuses et ont même permis la mise sur le marché de médicaments. Notons également que 80 % des protéines ne sont pas sécrétées ou ne sont pas présentes sur la membrane des cellules, mais à l'intérieur de ces dernières. Aussi les stratégies actuellement développées, qui consistent à utiliser de petites molécules organiques, ou d'imposants assemblages macromoléculaires incapables de traverser les membranes, ne peuvent être appliquées dans la grande majorité des cas. Une approche plus récente et de plus en plus développée, consiste à utiliser des objets de taille intermédiaire (0,5-5 kDa) pour cibler les IPP. Les foldamères, composés à la base de notre travail, appartiennent à cette dernière classe.

Après avoir rappelé les propriétés qui font des IPP des cibles à fort potentiel mais difficiles à atteindre, nous nous attacherons, dans un second temps, à présenter les multiples stratégies que les scientifiques ont développées afin de les inhiber, puis nous essaierons de montrer au lecteur que les foldamères en général, et les oligomères de quinolines en particulier, sont des molécules toutes indiquées pour l'inhibition d'IPP.



# 1. Interactions Protéine-Protéine (IPP)

## 1.1. Définition et classification des IPP

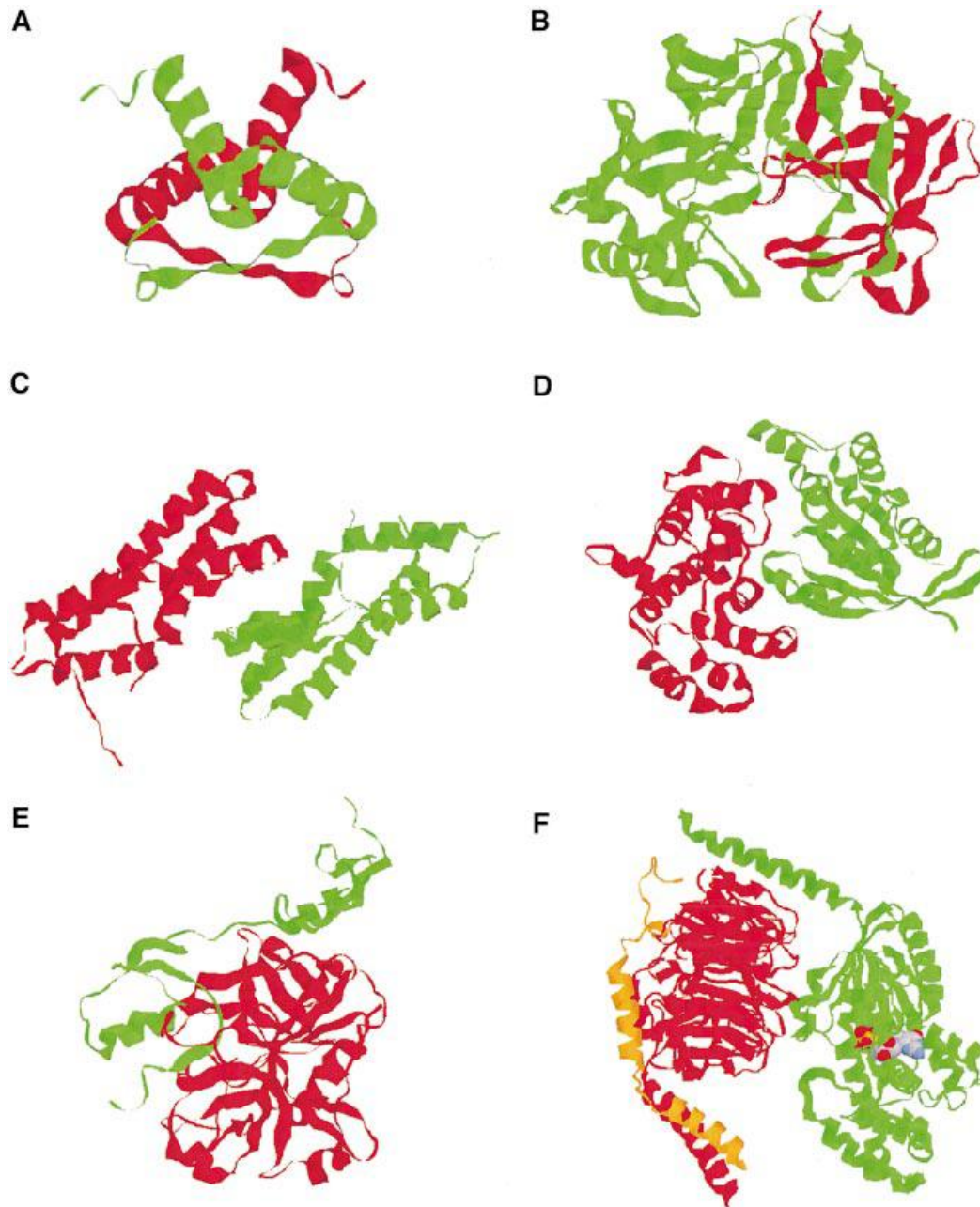
Les modalités de reconnaissance d'une protéine par une autre sont extrêmement diversifiées. Aussi leur classification est relativement difficile. Une des méthodes généralement utilisées consiste à distinguer les différents types de complexes en fonction de leur affinité et de leur fonction.

On peut différencier ces complexes suivant leur état multimérique et suivant la nature des partenaires (homo/hétéro dimères/oligomères), suivant leur propriétés structurales (taille de l'interface de quelques centaines à quelques milliers d'angströms carrés) ou suivant leur propriétés physiques (complexes obligatoires ou non et permanents ou transitoires) (Nooren et Thornton 2003).

Un complexe est considéré comme obligatoire lorsque les partenaires ne peuvent être séparés sans dénaturer les monomères (Jones et Thornton 1996), au contraire des complexes non-obligatoires. Les complexes sont dits permanents si les partenaires ne se dissocient jamais mais sont stables *in vivo* lorsqu'ils sont seuls.

Généralement les complexes obligatoires sont permanents (**Figure I.1. A et B**). Ils présentent des affinités très importantes ( $K_D$  de l'ordre du pM au nM). A l'opposé, les complexes non obligatoires (**Figure I.1. C et F**) sont souvent transitoires. Ils sont caractérisés par des affinités moins importantes ( $K_D$  de l'ordre du nM au  $\mu$ M). Les correspondances entre les aspects obligatoire et permanent, et entre les aspects non-obligatoire et transitoire ne sont pas toujours vraies. Il existe des complexes non-obligatoires mais permanents (Voir **Figure I.1. E**).

Ce sont les complexes non obligatoires et transitoires qui sont les plus représentés lorsque l'on s'intéresse aux processus de régulation cellulaire comme la transduction du signal ou les mécanismes d'oxydoréduction qui nécessitent une plasticité et une rapidité de réponse importantes (Janin 1995).

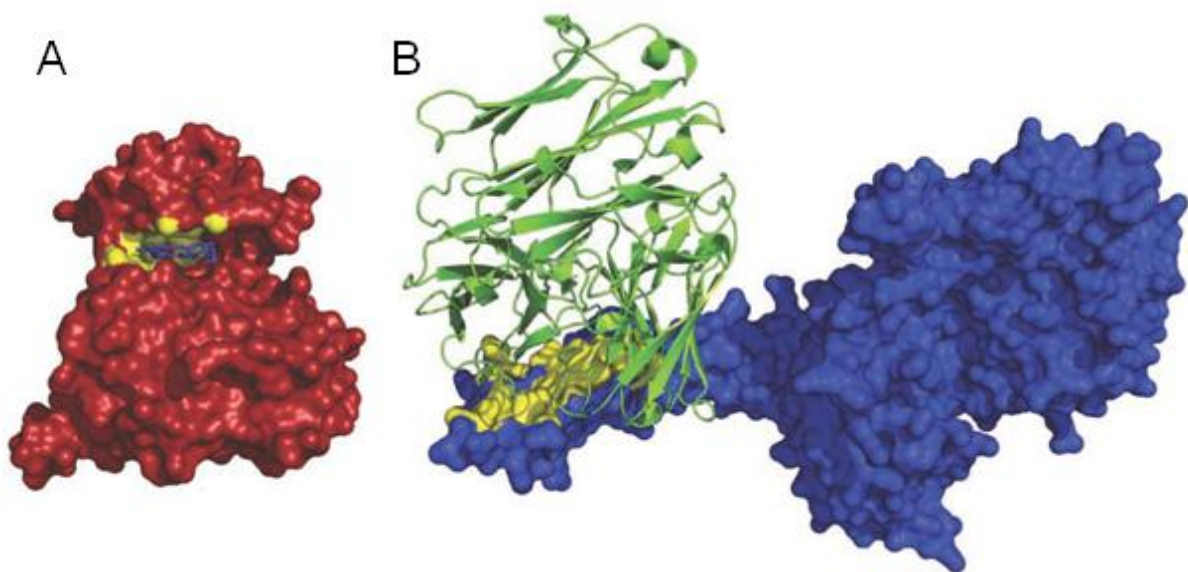


**Figure I.1. : Exemple de différents types d'IPP.**

**A : Homo-dimère obligatoire (répresseur Arc du phage P22). B : Hétéro-dimère obligatoire permanent (cathepsine D humaine). C : Homo-dimère non obligatoire transitoire (dimère de lysine d'ormeau). D : Hétéro-dimère non obligatoire (complexe de signalisation RhoA en vert et RhoGAP en rouge). E : Hétéro-dimère non obligatoire permanent (thrombine en rouge et l'inhibiteur rodniine en vert). F : Hétéro-trimère non obligatoire transitoire (protéine G bovine (sous unité  $G\alpha$  en vert) et le sous complexe  $G\beta\gamma$  en rouge et orange). (Nooren et Thornton2003).**

## 1.2. Structure et description des IPP

A l'inverse des enzymes qui sont dotées d'un site actif profond et bien localisé, permettant la liaison et l'interaction avec de petits ligands, la formation des IPP s'effectue par la reconnaissance de surfaces étendues et relativement planes (**Figure I.2.**). Les interactions hydrophobes, la complémentarité de forme et/ou de potentiel électrostatique qu'ils établissent entre leurs différents sites de reconnaissance, sans oublier la flexibilité des molécules impliquées (Jones et Thornton 1996), font de ces complexes des objets difficiles à étudier et à cibler.



**Figure I.2. : Comparaison d'une interaction enzyme-ligand et d'une IPP.**

**A : un inhibiteur (bleu) lié au domaine kinase ABL (rouge). B : l'herceptine (vert) liée au récepteur du facteur de croissance épidermique, EGFR (bleu). Les surfaces d'interaction sont représentées en jaune (Verdine et Walensky 2007).**

### 1.2.1. Surfaces de contact

La structure et l'affinité des IPP est en relation avec la fonction biologique sous-jacente, ainsi qu'avec l'environnement physiologique. Les interfaces offrent de ce fait une diversité très importante, tant en terme de taille, de forme que de composition. Les surfaces de contact impliquées dans les IPP sont relativement planes, sans relief et l'interaction entre les deux partenaires s'effectue par des interactions électrostatiques, des liaisons hydrogène, des

liaisons de van der Waals et par effet hydrophobe. Une complémentarité de forme et de charges des interfaces permet une spécificité élevée entre les sites d'interaction.

Les surfaces impliquées dans les complexes obligatoires et/ou permanents, présentent généralement des acides aminés hydrophobes, contrairement aux monomères impliqués dans les complexes non obligatoires et/ou transitoires qui présentent des surfaces plus polaires, voire même très chargées (Janin 1995).

Les surfaces d'interaction sont le plus souvent très importantes, entre 1 200 et 2 000 Å<sup>2</sup> comparées aux surfaces de contact entre enzyme et inhibiteur qui n'impliquent que des surfaces moyennes de 500 Å<sup>2</sup>. Dans une première approximation, il serait tentant de dire que la force, la spécificité de l'interaction, et le temps de demi-vie du complexe, sont totalement corrélés à la taille de l'interface. Dans la majorité des cas, cette règle s'applique. Néanmoins, dans d'autres cas, différents facteurs tels que l'environnement physiologique, la fonction biologique ou la nécessité de changements conformationnels pour induire la formation du complexe, entrent en jeu et modifient cette règle. Certaines interfaces sont petites (900-1 200 Å<sup>2</sup>) et correspondent à des complexes transitoires de stabilité faible, comme les complexes protéiques impliqués dans le transfert des électrons qui nécessitent de basses affinités ( $K_D$  de l'ordre du  $\mu\text{M}$ ) et des temps de demi-vie courts de l'ordre de la milliseconde (Janin *et al.* 2007). Les interactions antigène-anticorps suivent également cette règle. Leur surface d'interaction s'étend généralement de 1 200 à 2 300 Å<sup>2</sup>. Ce sont des complexes relativement stables avec des demi-vies de l'ordre de l'heure ou du jour et des constantes d'affinité de l'ordre du nanomolaire (Janin *et al.* 2007). Les interactions impliquant des interfaces de très grande taille 2 000-4 600 Å<sup>2</sup> se produisent dans le cas des protéines G ou avec les complexes intervenant dans les systèmes de transduction du signal. Ces complexes ont des temps de demi-vie beaucoup plus court de l'ordre de la milliseconde, et des constantes d'affinité de l'ordre du micromolaire (Janin 1995).

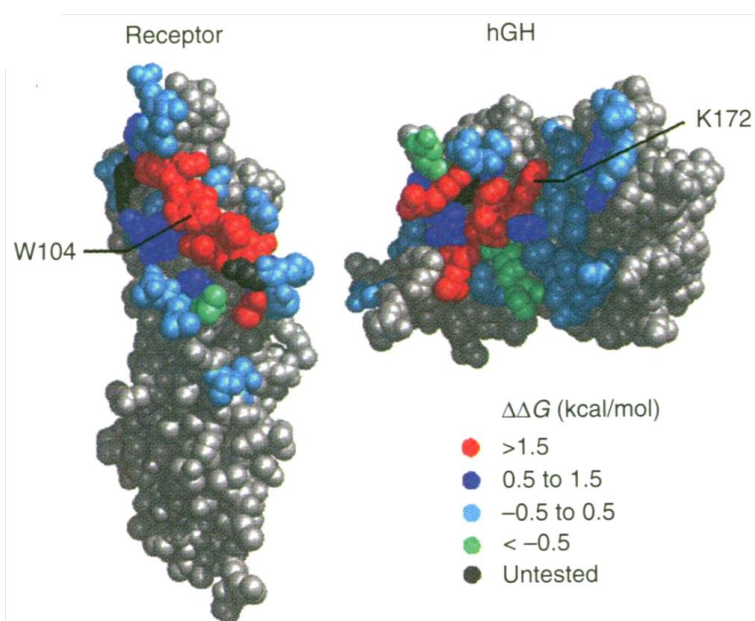
Cependant, l'énergie d'interaction entre ces surfaces étendues n'est, le plus souvent, pas répartie de façon uniforme. Certaines régions plus petites ( $\sim 500$  Å<sup>2</sup>), appelées *hot spot*, y contribuent plus fortement. Elles sont déterminantes pour la reconnaissance et l'affinité de liaison (Moreira *et al.* 2007).

### 1.2.2. Hot spots ou points de contact

Grâce aux études de mutagenèse dirigée, et à la caractérisation de ces mutants, certains résidus peuvent être identifiés comme étant les principaux acteurs responsables de

l'interaction (DeLano 2002). Les *hot spots* sont définis comme étant des résidus dont la mutation en alanine (*ala scanning*) entraîne une diminution de l'énergie d'interaction d'au moins 2,0 kcal/mol (Moreira *et al.* 2007). Le nombre moyen de résidus impliqués dans une IPP est de l'ordre de quelques dizaines, mais seuls quelques résidus (le plus souvent moins d'une dizaine) concentrent à eux seuls les trois quarts de l'énergie d'interaction.

Dans une étude réalisée par l'équipe de Wells (Clackson et Wells 1995), seuls 8 résidus parmi les 31 constituant l'interface du complexe que forme l'hormone de croissance humaine et son récepteur (hGH/hGHbp), sont responsables de 85 % de l'énergie de liaison (**Figure I.3.**).



**Figure I.3. : Complémentarité de l'interface du complexe hGH/hGHbp.**

Les résidus de l'interface sont indiqués selon leur contribution à l'énergie de liaison : en rouge, les résidus clés, responsables d'une grande partie de l'interaction (énergie de liaison > 1,5 kcal/mol) ; en bleu et vert, les autres résidus de l'interface (énergie de liaison < 1,5 kcal/mol) (Clackson et Wells 1995).

La composition en acides aminés de ces sites de reconnaissance diffère du reste de la surface de la protéine et présente, le plus souvent, une forte densité de résidus aromatiques, tyrosines et tryptophanes, ainsi que de résidus arginines (Tuncbag *et al.* 2009). Ces régions, pour la plupart situées au centre de l'interface, sont généralement entourées par des résidus périphériques, énergétiquement moins importants, qui ont pour rôle de favoriser les interactions en excluant les molécules d'eau.

L'identification des *hot spots* au sein d'une surface d'interaction entre deux partenaires peut être réalisée de plus en plus fidèlement. Le développement de méthodes de prédiction de ces résidus comme Robetta, KFC ou LDA, est également en plein essor (Tuncbag *et al.* 2009).

### 1.3. Détection et analyse des IPP

De nombreuses méthodes pour l'identification et la caractérisation des IPP se sont multipliées et sont actuellement disponibles. Ces principales méthodes peuvent être répertoriées selon quatre niveaux de résolution : atomique, moléculaire, complexe et cellulaire (Xenarios et Eisenberg 2001) (**Tableau I.1.**).

Tests	Niveaux d'observation			
	atomique	moléculaire	complexe	cellulaire
d'activité				•
biophysiques	•	•	•	
biologiques			•	•
Méthodes				
Rayons X et RMN	•	•		
Tests de compétition		•		
Retard sur gel		•		
ELISA		•	•	
Double-hybride		•	•	
Colonne d'affinité		•	•	
SPR		•	•	
Résonance paramagnétique électronique	•	•	•	
Chromatographie d'exclusion de taille		•	•	
Criblage par spectrométrie de masse		•	•	
Expérience de pontage chimique		•	•	
Co-immunoprécipitation		•	•	
Co-sédimentation		•	•	
Sédiment sur gradient de sucrose		•	•	
Co-purification			•	
Microscopie électronique			•	
Gel natif			•	
Immunoprécipitation			•	
Immunofluorescence				•
Expérience de FRET	•	•	•	•
Blocage par anticorps monoclonal			•	•
Extinction de gene ( <i>Knock-out</i> )				•
Co-expression transitoire				•

**Tableau I.1. : Tests et méthodes pour la détection des IPP.**

La caractérisation des interactions au niveau cellulaire permettra, entre autres, d'identifier les partenaires d'une interaction. Les méthodes d'étude au niveau du complexe permettent généralement de caractériser des complexes de grandes tailles (une à plusieurs dizaines de sous-unités) (taille, nombre de sous-unités, localisation grossière des sous-unités par rapport aux autres...). Les informations atomiques et moléculaires sont encore plus précises (d'après Xenarios et Eisenberg 2001).

Des méthodes telles que le double hybride, la co-immunoprécipitation, la chromatographie d'affinité couplée à la spectrométrie de masse, ou encore la méthode du FRET (*Fluorescence Resonance Energy Transfer*) ont permis la découverte d'un grand nombre d'IPP.

Une combinaison de plusieurs de ces techniques est souvent utile afin d'éliminer les faux-positifs et de confirmer les interactions. Une interaction peut, par exemple, s'effectuer de manière indirecte via la présence d'un partenaire commun, ou peut tout simplement être non physiologique. Les conditions spécifiques à certaines méthodes peuvent promouvoir des interactions qui n'existent pas *in vivo*. Une vérification de l'affinité des partenaires identifiés est donc nécessaire à travers des méthodes d'analyse complémentaires telles que la résonance plasmonique de surface (SPR) ou encore à travers l'utilisation d'analyses structurales comme la cristallographie ou la Résonance Magnétique Nucléaire (RMN).

La SPR permet d'accéder à l'affinité entre partenaires mais également de préciser les constantes d'association et de dissociation du complexe protéique. Mais il s'avère souvent essentiel de préciser au niveau atomique les interactions entre partenaires afin par exemple, de les inhiber dans un but thérapeutique. La cristallographie aux rayons X est, en ce sens, un outil primordial. Un grand nombre d'interactions décrites au sein de complexes protéine-protéine a été précisé par cette méthode comme le montre par exemple la résolution de la structure de complexes de grandes tailles, donc composés de nombreuses sous-unités, tels ceux de l'ADN polymérase (Zhang *et al.* 1999) ou du ribosome 70S contenant un ARN messager et un ARN de transfert (Yusupov *et al.* 2001). La RMN tient une place plus limitée, si l'on considère le nombre de complexes caractérisés par cette méthode. La raison essentielle réside dans les limitations que présente cette technique quant à la taille des objets étudiés (30 à 40 kDa le plus souvent) si l'on veut identifier avec précision l'origine intra ou intermoléculaire des signaux NOEs obtenus.

Grâce à de nouveaux outils informatiques, la prédiction des IPP à partir de la structure des différents partenaires est un domaine en pleine expansion. Avec l'explosion du nombre de structures résolues de protéines monomériques mais aussi de complexes, se construisent petit à petit diverses bibliothèques de modèles de plus en plus précis comme la PDB, la DIP (Salwinski *et al.* 2004) ou encore la BIND (Bader *et al.* 2003). Ainsi ces résultats seront-ils exploitables dans le cadre de la découverte de nouvelles molécules thérapeutiques, en ciblant spécifiquement les surfaces d'interaction potentielles mises en évidence. Le criblage de



ligands, la chimie combinatoire ou la conception rationnelle basée sur la structure, permettront d'élaborer de nouvelles molécules afin d'inhiber la formation de ces IPP.

Les IPP jouent un rôle majeur dans tous les processus cellulaires qu'il s'agisse de réplication, de métabolisme, de transport ou de signalisation. Chaque IPP est unique et présente un profil (taille, relief, polarité, ...) qui lui est propre. Les interfaces impliquées dans ces IPP sont souvent très étendues mais seuls quelques résidus plus ou moins répartis sur la surface sont responsables de la majeure partie de l'énergie d'interaction. Aussi il est évident que les petites molécules inhibitrices utilisées pour moduler les activités enzymatiques localisées dans de petits sites actifs profonds et particulièrement simples structurellement ne sont guère compatibles pour l'inhibition d'IPP.

## 2. Inhibition des IPP

---

Le dérèglement des interactions protéine-protéine, lié notamment à des mutations, est à l'origine de nombreuses pathologies, dont certains cancers (Verdine et Walensky 2007). L'enjeu qui consiste à mieux comprendre les phénomènes de reconnaissance moléculaire est primordial, non seulement pour décoder l'organisation supramoléculaire *in vivo*, mais également pour élucider les bases moléculaires de plusieurs maladies impliquant des altérations d'IPP (agrégations anormales de protéines avec Alzheimer par exemple) (Gonzalez et Kann 2012). Tout processus permettant de moduler, voire d'inhiber les interactions responsables de la formation de ces complexes, devrait donc permettre d'influer sur des événements cellulaires majeurs, à l'origine de diverses pathologies, et de développer de nouveaux agents thérapeutiques.

La modulation d'un processus par l'inhibition d'une IPP est susceptible de présenter plusieurs avantages par rapport à l'inhibition d'une enzyme. D'une part, l'extrême complémentarité des interfaces protéine-protéine offre la possibilité de développer des inhibiteurs hautement spécifiques, évitant ainsi l'inhibition d'autres cibles cellulaires. D'autre part, l'inhibition de telles interactions pourrait être d'une efficacité plus grande au niveau thérapeutique que l'utilisation d'un simple antagoniste pour une enzyme donnée dans la mesure où le risque d'échappement thérapeutique serait moindre (Veselovsky *et al.* 2002). L'argumentaire évoqué par l'auteur est le suivant. Dans le cas d'une enzyme, une simple



mutation au niveau du site actif (le résidu muté n'étant pas impliqué dans l'activité catalytique) peut rapidement générer une moindre affinité de l'antagoniste mais conserver son activité catalytique intacte. Dans le cas d'une IPP, l'échappement thérapeutique, donc la conservation du complexe induisant le processus à moduler, nécessiterait une mutation au niveau de l'une des interfaces mais également une mutation complémentaire à la surface du second partenaire, ce qui est statistiquement peu probable.

Chaque interface protéique étant relativement unique, la recherche d'inhibiteurs d'IPP diffère nettement d'une démarche plus classique qui repose sur l'inhibition d'une activité enzymatique par de petites molécules (généralement des criblages de vastes chimiothèques). Dans le cas de larges surfaces, de nouvelles approches (de synthèse, biologique, ou de conception rationnelle) nécessitent d'être échafaudées pour chaque nouvelle interaction étudiée. De plus les tests biologiques d'affinité d'un inhibiteur d'IPP ne relèvent souvent pas d'une activité enzymatique. Il est nécessaire d'utiliser des méthodes plus pointues pour caractériser l'efficacité d'un inhibiteur comme la titration calorimétrique isotherme (ITC) ou la résonance plasmonique de surface (SPR).

La conception d'un inhibiteur d'une IPP, repose sur l'agencement spatial de groupements chimiques, biologiques ou non (chargés, polaires, hydrophobes, ...), dicté par une molécule support (elle aussi chimique ou biologique). Cet agencement doit être complémentaire d'un des partenaires pour assister à l'inhibition de l'interaction. Cette complémentarité équivaut généralement à mimer le(s) domaine(s) ou les *hot spots* responsables de l'interaction du second partenaire. Dans cette partie nous montrerons que la nature met à notre disposition des molécules supports (oligonucléotides, anticorps, protéines miniatures, ...) portant des groupements chimiques modulables (par criblage ou par conception rationnelle). Nous nous attacherons également à convaincre le lecteur que des molécules de synthèse peuvent être utilisées comme molécules supports de manière très efficace.

## 2.1. Des objets biologiques pour inhiber les IPP

### 2.1.1. Oligonucléotides fonctionnels

La potentialité des oligonucléotides à lier des ligands spécifiques est largement utilisée pour le développement de réactifs pharmaceutiques et diagnostiques. Depuis 1990, avec la découverte d'une méthode combinatoire de sélection *in vitro* (SELEX Stoltenburg *et al.*

2007), de nombreux aptamères ont été identifiés. Ces composés peuvent également être modifiés, par greffage de divers groupements, les rendant plus stables *in vivo* (Osborne et Ellington 1997). Ils peuvent être sélectionnés contre des cibles de natures très diverses (petites molécules organiques, peptides, protéines, acides nucléiques, cellules) et sont en général caractérisés par des affinités élevées pour leur cible et par une spécificité remarquable.

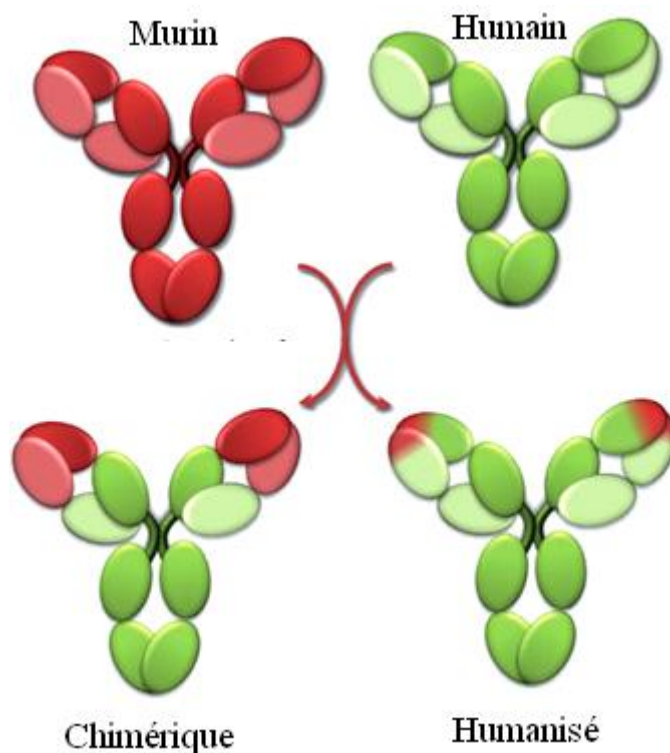
Le potentiel thérapeutique d'inhibition des interactions protéine-protéine de ces aptamères repose sur leur taille importante et leur large surface leur permettant d'interagir de manière très spécifique avec les surfaces biologiques (Osborne et Ellington 1997). Par exemple, des acides nucléiques résistants aux nucléases, et qui présentent une forte affinité ( $K_i$  de l'ordre du pM) pour le facteur de croissance endothélial vasculaire (VEGF), ont été identifiés à partir de bibliothèques d'aptamères ARN (Jellinek *et al.* 1994).

La taille de ces objets constitue à la fois un avantage (stabilité d'interaction) et des inconvénients (difficultés à traverser les membranes cellulaires et nucléaires, biodisponibilité faible). Les aptamères ne peuvent pas être délivrés oralement et doivent être injectés directement au patient, ou administrés à l'aide de transporteurs, ou par encapsulation dans des liposomes (Osborne et Ellington 1997). Même si certaines modifications chimiques permettent une biodisponibilité améliorée (Osborne et Ellington 1997), leur coût de production (essentiellement dû à leur taille) constitue une limitation à l'utilisation de ces molécules qui restent cependant moins volumineuses que les anticorps (~ 150 kDa).

### 2.1.2. Anticorps monoclonaux

Encore récemment, l'idée d'utiliser un anticorps comme objet thérapeutique *in vivo* était considéré comme ambitieuse. Le développement d'une nouvelle génération d'anticorps recombinants monoclonaux a permis d'entrevoir un potentiel clinique considérable touchant à différents domaines de la médecine : oncologie, cardiologie, inflammation, rhumatologie, transplantation, dermatologie...

Les anticorps monoclonaux recombinants sont des molécules artificielles qui interagissent de façon très spécifique en exposant des chaînes latérales spatialement définies et complémentaires d'un antigène. Les anticorps monoclonaux chimériques ou humanisés sont obtenus en échangeant une partie d'un anticorps humain par une partie d'un anticorps de souris. Seule la partie de l'anticorps de la souris impliquée dans la liaison de l'antigène ciblé, est transférée (**Figure I.4.**).



**Figure I.4. : Anticorps chimérique et humanisé.**

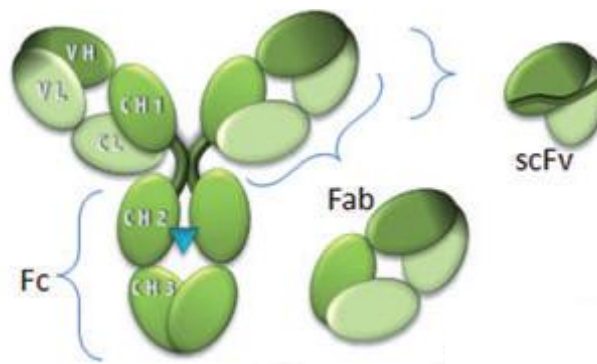
Les séquences de la souris sont représentées en rouge et celles de l'homme sont en vert. Les anticorps monoclonaux chimériques sont humains à environ 70 %, et les anticorps monoclonaux humanisés le sont à 90 %, ce qui permet de limiter les risques de rejet et d'accroître leur efficacité (Chames *et al.* 2009).

Les deux familles précédentes constituent les anticorps thérapeutiques. Ces derniers possèdent d'excellentes propriétés facilitant l'inhibition des interactions protéine-protéine : ils présentent une haute spécificité pour leur cible, sont peu toxiques et particulièrement stables en milieu physiologique (leur demi-vie pouvant atteindre dans certains cas une vingtaine de jours) (Chames *et al.* 2009). Leur taille importante ( $\sim 150$  kDa) leur permet d'interagir avec les larges interfaces des complexes protéine-protéine. Plus d'une vingtaine d'anticorps monoclonaux, murins, chimériques ou humanisés, sont actuellement commercialisés et au moins 400 d'entre eux sont testés dans le cadre d'essais cliniques de phase I à III (Stockwin et Holmes 2003). Depuis 1998, l'anticorps monoclonal humanisé, Trastuzumab (Herceptine<sup>TM</sup>, Genentech), dirigé contre l'homo-dimérisation du complexe Her2/Neu impliqué dans le mécanisme de prolifération cellulaire, est proposé dans le traitement des cancers du sein (Stockwin et Holmes 2003). Le Basiliximab (Simulect<sup>TM</sup>, Novartis), anticorps monoclonal humanisé qui bloque l'interaction entre l'interleukine 2 (IL-2) et son récepteur (IL-2R), représente un exemple concret de l'inhibition d'interactions protéine-protéine en

immunothérapie. Il permet une réduction des risques de rejet chez les transplantés du rein en inhibant IL-2R, diminuant l'activation et la prolifération des lymphocytes T, responsables des rejets de transplants (Stockwin et Holmes 2003).

Néanmoins, les difficultés liées à la production et au coût de ces anticorps sont des facteurs limitant leur utilisation. Notons également que leur poids moléculaire important ne leur permet pas de passer les barrières cellulaires ou veineuses et sont donc très peu biodisponibles. Ils sont de ce fait réservés à des cibles extracellulaires et impliquent un mode d'administration plus difficile pour les patients (injection par voie intraveineuse ou intramusculaire).

D'autres approches basées sur des anticorps monoclonaux sont aujourd'hui activement explorées. L'utilisation de fragments d'anticorps est notamment développée pour bloquer ou moduler les fonctions de protéines intracellulaires (**Figure I.5.**). Les coûts de production sont moins élevés (production possible dans des systèmes microbiens tels que *E. coli* ou chez la levure), et leur taille (~ 50 kDa) est mieux adaptée à la pénétration des tissus. Ces entités conservent leur affinité de liaison et leur spécificité, permettant ainsi de développer leur utilisation pour de nombreuses nouvelles cibles thérapeutiques.



**Figure I.5. :** Exemple de fragments d'anticorps.

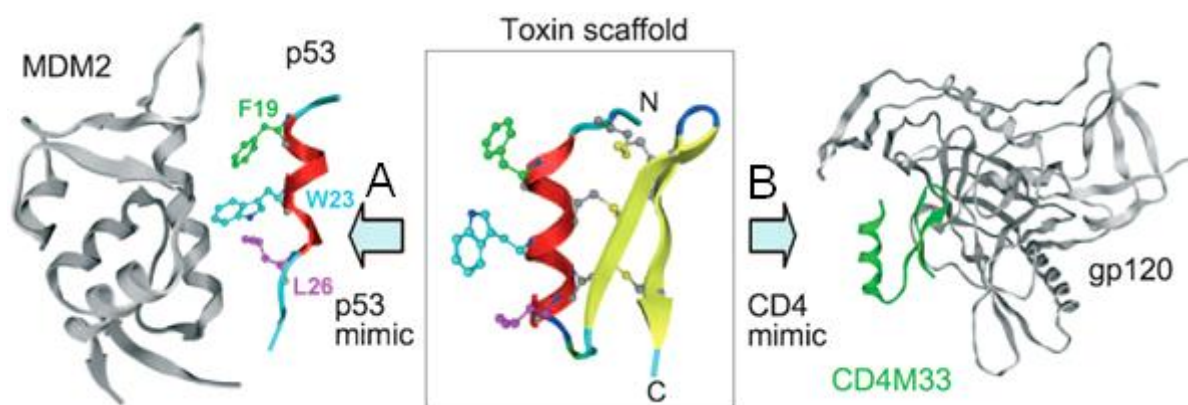
L'anticorps monoclonal entier, composé de sa partie Fc (*crystallizable fragment*) et de sa partie Fab (*antigen binding fragment*). La partie variable responsable de la reconnaissance est composée des domaines VL et VH (*Light and Heavy Variable domain*). Les fragments Fab et scFv (*single chain variable fragment*) sont utilisés en radioimmunothérapie, ou encore en imagerie (Stockwin et Holmes 2003).

### 2.1.3. Protéines miniatures

Une autre manière d'élaborer des objets antagonistes des IPP consiste à générer des mimes d'épitopes de protéines (Robinson 2009). Comme nous l'avons détaillé précédemment

une molécule support, jouant le rôle d'échafaudage conformationnellement contraint (adoptant une structure spatiale très stable), expose des fonctions chimiques spatialement définies. Le choix de la protéine support (si c'est une protéine) est donc un paramètre clé pour la conception de mimes d'épitopes car il permet la disposition optimale des résidus importants pour l'interaction. Une fois la protéine support choisie, le transfert des résidus de l'épitope sur la protéine permet d'obtenir un inhibiteur qui fera l'objet d'une optimisation. Beaucoup de groupes de recherche utilisent ainsi des protéines ou des fragments de protéines particulièrement stables et naturellement existants, comme molécule support. La particularité de cette approche est qu'elle repose sur le mimétisme d'éléments de structures secondaires bien définis.

La toxine (BmBKTx1) de 31 acides aminés, issue du venin du scorpion asiatique *Buthus martensi*, a été utilisée comme molécule support. Le peptide comprend une hélice  $\alpha$  ainsi qu'une boucle en épingle à cheveux positionnée entre deux brins  $\beta$ . L'épitope impliqué dans la reconnaissance de p53 par MDM2 est une hélice  $\alpha$ . Cette dernière possède les résidus Phe19, Trp23 et Leu26 déterminants pour l'interaction. Ces résidus ont été transférés sur le segment hélicoïdal de la toxine de scorpion par mutation, conférant au mime une grande affinité ( $K_D = 790$  nM) pour la protéine MDM2 (Li *et al.* 2008) (**Figure I.6.**).



**Figure I.6. : Utilisation d'un échafaudage basé sur une toxine de scorpion.**

**A :** Transfert de l'épitope de p53 sur l'hélice  $\alpha$  de la toxine pour bloquer l'interaction p53/MDM2 (PDB : 1YCR). **B :** Transfert de l'épitope en boucle en épingle à cheveux  $\beta$  de CD4 sur la toxine pour bloquer l'interaction CD4/gp120 (PDB : 1YYL) (Robinson 2009).

Le motif en épingle à cheveux de cette toxine a également été utilisé pour générer des mimes de l'épitope de CD4 qui interagit avec la glycoprotéine gp120 du virus de l'immunodéficience humaine de type 1 (VIH-1) (Martin *et al.* 2003). Cet épitope a donc été

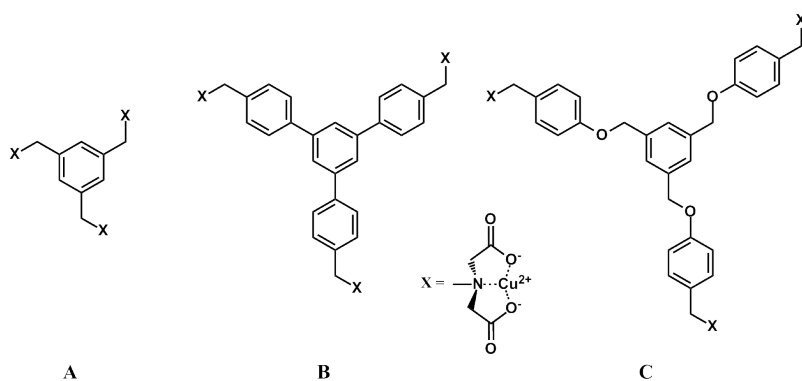
transféré sur la toxine du scorpion par mutation, permettant, après optimisation, d'obtenir des mimes qui lient la protéine gp120 avec une très forte affinité et inhibe ainsi l'internalisation du VIH-1 dans les cellules (**Figure I.6.**).

## 2.2. Des objets de synthèse pour inhiber les IPP

### 2.2.1. Systèmes à base de métaux

La surface d'une protéine offre un large choix, nous l'avons déjà mentionné, de groupes chargés, polaires, hydrophobes, aliphatiques et aromatiques qui peuvent être ciblés par des groupements chimiques pour la reconnaissance de surface. Certains métaux de transition ( $\text{Cu}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ ) peuvent se coordonner très facilement aux fonctions imidazole des résidus histidines exposées à la surface d'une protéine. Leur utilisation pour la reconnaissance de surfaces de protéine induit des liaisons covalentes de coordination métal-ligand spécifiques, plus fortes énergétiquement que toutes les autres interactions non-covalentes existant généralement entre partenaires d'un complexe. Cette stratégie permet donc d'obtenir une affinité et une sélectivité plus importantes tout en limitant le nombre d'interactions à générer avec la surface protéique.

Fazal *et al.* (2001) ont décrit la conception et la synthèse de molécules à base de métaux de transition visant les six résidus histidines présents à la surface de l'anhydrase carbonique. L'affinité de liaison des molécules pour la protéine a ensuite été étudiée en observant les changements des propriétés spectroscopiques du métal. Les ligands possèdent trois motifs  $\text{Cu}^{2+}$ -iminodiacétates séparés par des espaceurs aromatiques de différentes longueurs (**Figure I.7.**).



**Figure I.7. : Dérivés  $\text{Cu}^{2+}$ -iminodiacétates, utilisés pour reconnaître la surface de l'anhydrase carbonique.**

Sur les trois composés synthétisés, le deuxième composé (**Figure I.7.B**) est celui qui présente la meilleure constante d'affinité ( $K_D = 3,3 \mu\text{M}$ ). La sélectivité de la molécule pour l'anhydrase carbonique a été démontrée en mesurant une affinité plus faible avec l'albumine de l'œuf de poulet ( $K_D = 1 \text{ mM}$ ) qui possède également six résidus histidines à sa surface ainsi qu'avec la myoglobine ( $K_D = 50 \mu\text{M}$ ) qui en présente sept. Un espaceur plus court (**Figure I.7.A**) ou plus long (**Figure I.7.C**) induit une affinité moindre. La suppression des ions  $\text{Cu}^{2+}$  inhibe toute affinité de ces composés pour la protéine.

Cet exemple souligne l'importance de l'organisation spatiale des dérivés  $\text{Cu}^{2+}$  par rapport à la surface de la protéine et donc du choix de la molécule support.

### 2.2.2. Mimes de feuillets $\beta$

Les feuillets  $\beta$  constituent une grande famille de structures secondaires présentes dans les protéines. Ils représentent à eux seuls 30 % des éléments de structure. Cependant, peu d'exemples de structures synthétiques mimant les feuillets  $\beta$ , dans le but de moduler les interactions protéine-protéine, existent à ce jour. Le feuillet  $\beta$  est certainement l'élément structural le plus difficile à mimer car il implique des liaisons hydrogène entre des résidus très éloignés. C'est pour cela que le développement des structures se repliant ou mimant un feuillet  $\beta$  n'en est qu'à ses débuts.

Un exemple de molécule mimant un épitope présent sur un feuillet  $\beta$  pour inhiber une interaction protéine-protéine est donné, avec l'inhibition de la formation du complexe ICAM-1/LFA-1. L'interaction de la protéine d'adhérence ICAM-1 avec l'intégrine LFA-1 joue un rôle important dans le traitement des maladies auto-immunes, provoquant l'adhésion, la migration et la prolifération des lymphocytes (Gadek *et al.* 2002). Un épitope discontinu contenant les résidus Glu34, Lys39, Met64, Tyr66, Asn68 et Gln73, et localisé sur la face opposée à un feuillet  $\beta$  du domaine d'ICAM-1, a été identifié comme responsable de l'interaction avec LFA-1 (**Figure I.8.1**). L'équipe de recherche menée par Gadek (Gadek *et al.* 2002) a présenté une approche dans laquelle un inhibiteur (**Figure I.8.2A**) de cette interaction, sélectionné par criblage, a subi le transfert de l'épitope porté par ICAM-1, dans le but d'améliorer sa spécificité et son efficacité biologique. Ces modifications ont mené au composé **2B**, antagoniste de ICAM-1 par sa liaison à LFA-1 avec une très haute affinité ( $\text{IC}_{50} = 1,4 \text{ nM}$ ).

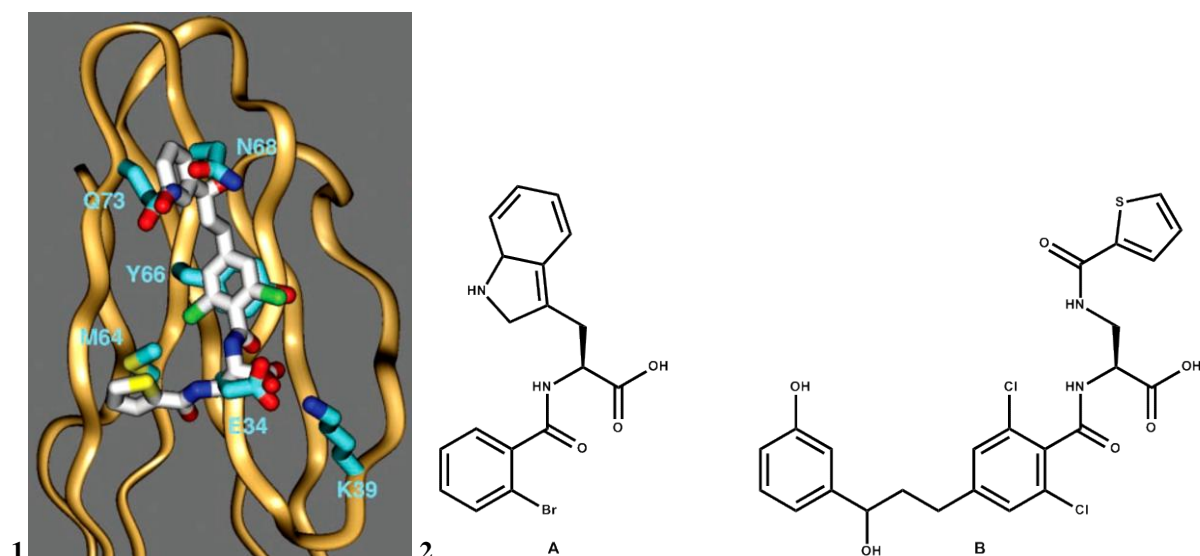


Figure I.8. : Inhibition du complexe ICAM-1/LFA-1.

1 : Superposition du composé 1B en blanc sur la structure du domaine d'ICAM-1 responsable de l'interaction avec LFA-1. Les résidus représentés en bâtonnets bleus font partie de l'épitope reconnu par LFA-1. 2A : Molécule à partir de laquelle les améliorations ont été faites pour conduire à la molécule 2B. (Gadek *et al.* 2002).

### 2.2.3. Mimes d'hélices $\alpha$

Bon nombre d'interactions protéine-protéine utilisent l'hélice  $\alpha$  comme élément clé dans la reconnaissance de surface. L'intérêt d'élaborer des mimes de telles structures est donc considérable et ces protéomimétiques constituent des candidats importants. Ils peuvent servir d'échafaudage pour une fonctionnalisation rationnelle et être directement utilisés comme antagonistes des interactions protéine-protéine.

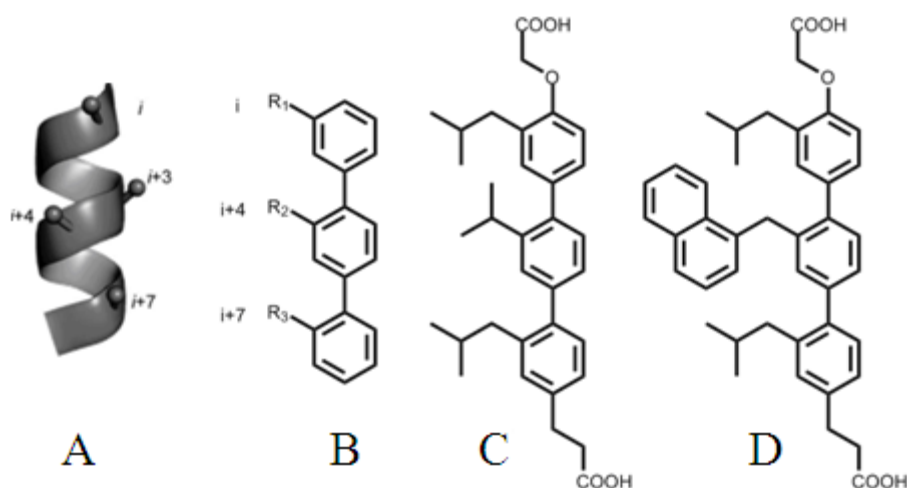
La structure hélicoïdale de ces composés est obtenue à partir de la répétition de motifs structuraux susceptibles d'apporter une contrainte directionnelle (Saraogi et Hamilton 2008). Comme dans les hélices  $\alpha$ , cette structure peut être décorée en utilisant diverses chaînes latérales susceptibles d'interagir avec la cible protéique.

Des exemples récents comprennent la conception rationnelle de structures non peptidiques, comme des dérivés terphényles (Yin et Hamilton 2005), qui distribuent tout le long d'une direction d'élongation donnée différents radicaux dont la périodicité respecte celle des chaînes latérales des résidus  $i$ ,  $i+3$  ou  $i+4$  et  $i+7$  d'une conformation hélicoïdale (**Figure I.9.A et B**). La fonctionnalisation appropriée de ces structures peut inhiber certains complexes protéine-protéine. Ainsi, le greffage de résidus hydrophobes tels que la leucine et la valine



permet-il, à une structure terphényle (**Figure I.9.C**) de mimer le domaine hélicoïdal de la protéine hexamérique gp41, responsable de la fusion du VIH-1 avec les cellules humaines, et donc d'inhiber son auto-assemblage (Yin et Hamilton 2005). Des structures contrôles, où le squelette ne présente que deux unités et non trois, ou ne possède plus de résidus hydrophobes, perdent tout pouvoir inhibiteur, démontrant ainsi le rôle conjugué joué par les 3 groupements hydrophobes dans le processus de reconnaissance de surface.

Les mêmes auteurs ont développé des antagonistes de l'interaction Bak-BH3/Bcl-x<sub>L</sub> en insérant une chaîne 1-naphthylméthylène sur la deuxième unité du motif terphényle (**Figure I.9.D**). Une telle modification permet en effet de mimer plus fidèlement la distribution des résidus clés le long de l'hélice  $\alpha$  du domaine BH3 de Bak, qui présente une grande affinité pour sa cible.



**Figure I.9. : Conception rationnelle de mimes d'hélices  $\alpha$  basés sur une structure terphényle.**

**A :** Représentation d'une hélice  $\alpha$  modèle avec ces résidus  $i$ ,  $i+3$ ,  $i+4$  et  $i+7$  mis en avant. **B :** le squelette terphényle avec les chaînes latérales  $R_1$ ,  $R_2$ , et  $R_3$  retrouvées dans les mêmes positions que les résidus  $i$ ,  $i+4$  et  $i+7$ . **C :** Formule de l'inhibiteur de l'auto-assemblage de gp41. **D :** Antagoniste de Bcl-x<sub>L</sub>.

Ces exemples démontrent l'adaptabilité du motif terphényle, capable de sélectivement cibler et modifier différentes interactions protéine-protéine.

La communauté scientifique et plus particulièrement les compagnies pharmaceutiques se sont évertuées à trouver de nouvelles classes de molécules pour inhiber ces cibles thérapeutiques prometteuses que sont les IPP. Certaines approches font appel à des molécules biologiques (anticorps, protéines miniatures), lorsque d'autres utilisent des objets de synthèse.

Le point commun à chacune de ces techniques réside dans l'utilisation d'une molécule support permettant d'exposer et de maintenir des groupements chimiques dans une conformation spatiale particulière : un épitope. Les propriétés de biodisponibilité, de non-toxicité, de perméabilité membranaire, etc, sont principalement conférées par la molécule support.

### 3. Les foldamères, des molécules qui se plient en quatre

Les propriétés fonctionnelles des macromolécules biologiques, telles que la reconnaissance moléculaire, le transfert d'informations ou encore la catalyse sont très souvent fortement corrélées au type de repliement adopté par la séquence polypeptidique. Ce repliement résulte de la combinaison d'interactions généralement non covalentes (hydrophobes, électrostatiques, induites par liaisons hydrogène ou par empilement de groupements aromatiques) et parfois covalentes (ponts disulfures, liaisons de coordination par l'intermédiaire d'un métal). Du fait de ces interactions, les polymères naturels (protéines) adoptent des structures secondaires (hélices, feuillet, coude, ...).

Ainsi, les chimistes s'investissent-ils activement dans la conception *de novo* de molécules dites biomimétiques, c'est-à-dire de taille et de structuration identiques à celles des macromolécules biologiques. Ces composés artificiels, appelés foldamères miment la structure et *a priori* les fonctions des peptides (Muller *et al.* 2009), tout en améliorant leurs propriétés natives. De récents développements ont montré que les structures secondaires des protéines, telles que les hélices  $\alpha$  et les feuillets  $\beta$ , pouvaient être mimées par ces foldamères (Appella *et al.* 1997, Vasudev *et al.* 2009).

L'utilisation de molécules synthétiques de taille intermédiaire (0,5-5 kDa), tels les foldamères, permettrait de compléter les stratégies actuellement utilisées pour cibler les interactions protéine-protéine. Ces composés, nous y reviendrons, ont des tailles comparables à celles de petites protéines. Ils présentent donc *a priori* de potentielles surfaces d'interaction directement comparables à celles impliquées dans les IPP. De par leurs dimensions, ils se rapprocheraient de certains objets biologiques, tels les fragments d'anticorps, dont nous avons

pu montrer l'intérêt dans la reconnaissance de surfaces protéiques. De tels composés ont d'autres atouts :

- Contrairement aux protéines ou peptides dont les règles de repliement sont complexes (de par la diversité des acides aminés présents dans la séquence), le repliement de tels oligomères est plus facile à prévoir. Ils sont en effet le plus souvent constitués d'un seul type de monomère (ou un nombre restant très limité) dont la répétition impose une géométrie d'interactions constante le long de la chaîne, directement responsable du repliement observé.
- Les monomères synthétiques utilisés pour les concevoir leur confèrent une résistance accrue à la dégradation enzymatique et peuvent potentiellement remplacer de manière avantageuse des  $\alpha$ -peptides ou des oligonucléotides comme substance active.
- Ces composés sont également hautement fonctionnalisables par des chaînes latérales qui, contraintes spatialement par le squelette du foldamère (molécule support), pourront leur conférer spécificité et affinité vis-à-vis d'une surface protéique.
- Parce que la synthèse de ces composés est généralement aisée, efficace et convergente, ces molécules sont compétitives, en terme de coût de production, avec de gros objets biologiques (anticorps, oligonucléotides).

Ces composés apparaissent donc comme des candidats appropriés pour se lier aux surfaces hautement structurées des protéines.

### 3.1. Différents squelettes ; différentes structures

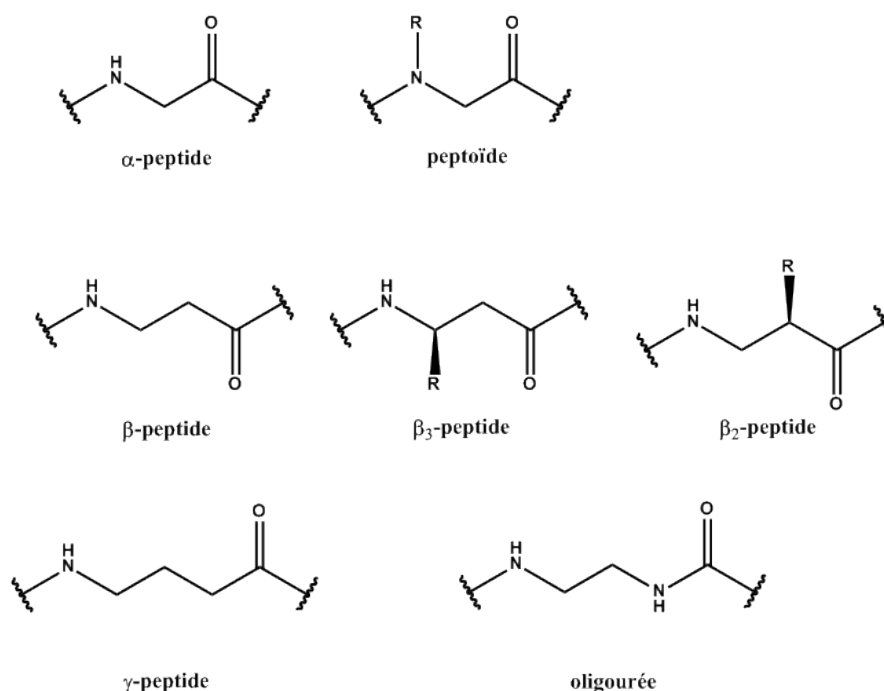
Le terme « foldamère » a, pour la première fois, été défini par Samuel Gellman, en 1998, comme étant « tout oligomère ayant une forte tendance à adopter une conformation compacte bien définie et prévisible en solution » (Gellman 1998). Depuis une quinzaine d'années, la synthèse et l'étude structurale de ces nouvelles molécules, destinées à mimer les structures secondaires des peptides et des protéines, font l'objet d'une recherche intensive. Aujourd'hui, la diversité des monomères utilisés, des structures obtenues, accroît de manière spectaculaire le potentiel de ces molécules à interagir avec des objets biologiques.

Les foldamères sont, comme les protéines, des oligomères constitués d'unités ou monomères de base dont la nature de la chaîne latérale peut varier. Il existe différentes manières d'enchaîner ces monomères : l'utilisation de la liaison peptidique ou liaison amide est très souvent utilisée, mais l'utilisation de la liaison urée ( $-\text{NH}(\text{CO})\text{NH}-$ ) (**Figure I.10.**) permet d'apporter une variation du squelette du foldamère.

Il existe deux grandes familles de foldamères. La première, dite bio-inspirée, consiste à modifier les systèmes biologiques existants en y apportant des améliorations (stabilité, perméabilité...). Cette famille inclut essentiellement les foldamères aliphatiques. La seconde famille tente de mimer les structures secondaires des macromolécules biologiques en utilisant des squelettes innovants, qui n'existent pas dans la nature et donc à concevoir *de novo* de nouveaux motifs de reconnaissance. Ce sont les foldamères aromatiques.

### 3.1.1. Les foldamères aliphatiques

Beaucoup d'efforts se sont portés sur l'amélioration des macromolécules biologiques en élaborant des foldamères à partir de  $\beta$ -peptides, de  $\gamma$ -peptides, de  $\delta$ -peptides et autres composés dérivés (**Figure I.10.**).

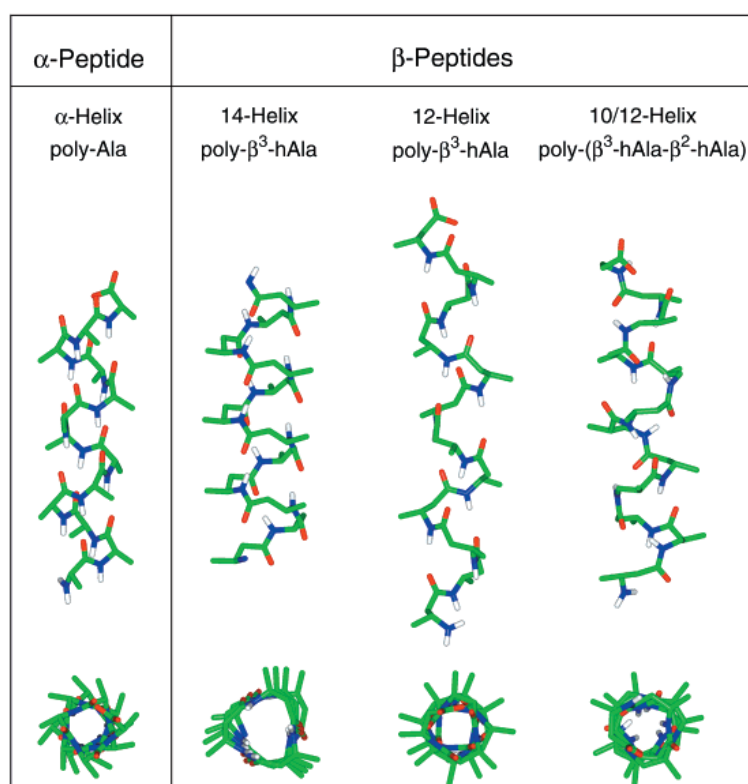


**Figure I.10. : Exemples de différents monomères utilisés dans la conception de foldamères aliphatiques.**

Les  $\beta$ -peptides, sont formés d'acides aminés  $\beta$ , assemblés de manière linéaire et capables de se replier en des structures secondaires stables grâce à des liaisons hydrogène intramoléculaires. Ils se différencient des  $\alpha$ -peptides par la présence d'un atome de carbone supplémentaire qui leur confère des propriétés pharmacodynamiques favorables et une bonne perméabilité membranaire (Seebach et Gardiner 2008). Une des propriétés importantes de ces objets est leur résistance à la protéolyse et à l'hydrolyse (Disney *et al.* 2005), aux enzymes

métaboliques des mammifères, des insectes et des plantes (Lind *et al.* 2004), ainsi qu'aux enzymes des microorganismes (Heck *et al.* 2006 et Seebach *et al.* 2004) ; de plus il est très rare de leur attribuer des activités antibiotiques ou hémolytiques (Arvidsson *et al.* 2005).

Seebach (Seebach et Gardiner 2008) et Gellman (Appella *et al.* 1997) ont été les pionniers du développement des  $\beta$ -peptides et ont très vite démontré que cette classe de protéomimétiques peut adopter une variété de conformations hélicoïdales bien distinctes (**Figure I.11.**) ou de conformations en motif d'épingle à cheveux. Les différences de conformation observées sont fonction du choix de l'unité constitutive, de leur sens d'hélicité ainsi que du choix des substituants présents sur ces unités. Elles conduisent à l'observation d'hélices 10, d'hélices 10/12, d'hélices 12, d'hélices 14 principalement.

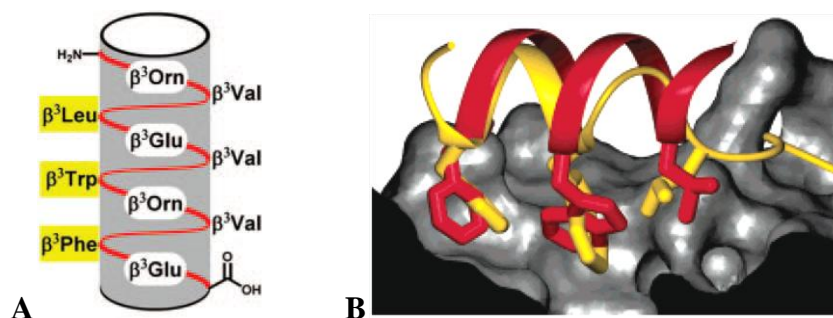


**Figure I.11. : Structure des hélices  $\alpha$ , 14, 12, 10/12 construites à partir de résidus  $\alpha$  ou  $\beta$ .**

Les atomes d'hydrogène ont été supprimés pour plus de clarté, à l'exception de ceux portés par les atomes d'azotes (blanc). Les atomes de carbone, d'azote et d'oxygène sont représentés respectivement en vert, bleu et rouge (Cheng *et al.* 2001).

Des inhibiteurs spécifiques des IPP, basés sur des  $\beta$ -peptides ainsi que sur des chimères  $\alpha/\beta$ -peptides, peuvent alors être développés si les surfaces des protéines cibles impliquent une hélice  $\alpha$  comme site de reconnaissance. C'est ainsi que l'équipe de Seebach

(Werder *et al.* 1999) a utilisé un  $\beta$ -peptide de type hélice 14 pour mimer l'hélice  $\alpha$  des apolipoprotéines impliquées dans l'absorption et le transport du cholestérol de l'intestin grêle jusqu'aux systèmes lymphatique et vasculaire. Ces molécules se révélèrent être de très bons inhibiteurs de l'absorption du cholestérol. L'équipe de Schepartz (Kritzer *et al.* 2005 a et b) a également développé des décimères se repliant en hélice 14 pour cibler les sites de reconnaissance de p53 sur hMD2 (**Figure I.12.**).



**Figure I.12. : Inhibition du complexe p53/hMD2 par un  $\beta$  peptide.**

**A :** Structure schématique du décimère. **B :** Superposition de la structure du décimère (ruban et chaînes latérales en rouge) sur la structure de l'hélice du domaine d'activation du peptide p53 (ruban et chaînes latérales en jaune) en interaction avec hMD2 (surface grise). Seules les chaînes latérales impliquées dans l'interaction sont représentées (Kritzer 2005)

Des travaux très récents ont montré que l'utilisation de  $\beta$ -peptides hybrides, formés d'une alternance d'acides  $\alpha$ - et  $\beta$ -aminés que l'on appelle  $\alpha/\beta$ -peptides était pertinente vis-à-vis de l'inhibition d'IPP (Johnson et Gellman 2013, Smith *et al.* 2013). En effet cette équipe a développé des molécules qui miment le domaine de reconnaissance BH3 de la protéine pro-apoptotique Bak, dans le but d'inhiber la formation du complexe Bcl-xL/BH3-Bak. Pour cela une molécule alternant des résidus  $\alpha$  et des résidus  $\beta$ , suivis d'un segment de six résidus  $\alpha$  (**Figure I.13.A**) a été synthétisée. Elle présente une affinité très élevée (1 nM) pour le site de reconnaissance de Bcl-xL. Le peptide  $\alpha/\beta+\alpha$  a été co-cristallisé en présence de Bcl-xL (PDB = 3FDM, **Figure II.3.B et C**) et des études ont montré que ce type de composé bloque l'interaction entre Bcl-xL et sa protéine complémentaire dans un lysat cellulaire (Sadowsky *et al.* 2007).

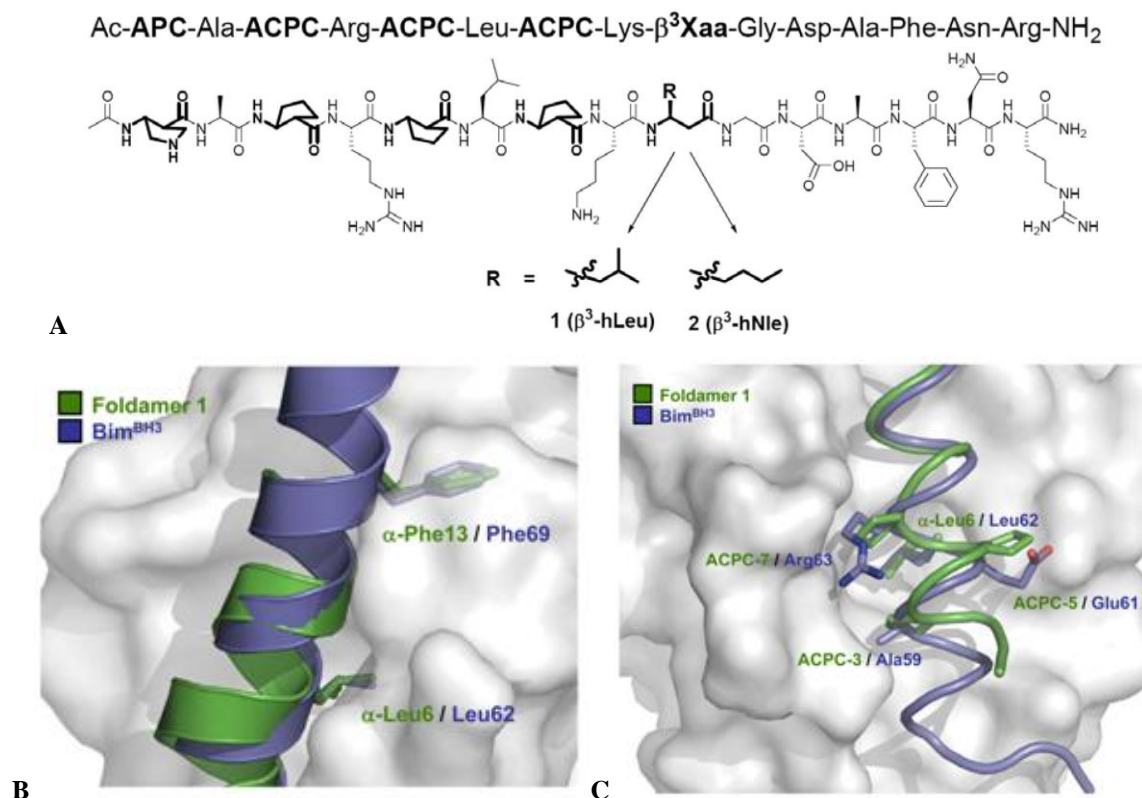


Figure I.13. : Inhibition du complexe Bcl-xL/BH3-Bak par un  $\alpha/\beta$ -peptide (Lee *et al.* 2009).

A : Séquence du pentadécamère ayant démontré une activité biologique. B et C : Comparaison du domaine BH3 sous la forme du peptide Bim<sup>BH3</sup> et du foldamère mime. B : Mise en évidence de la superposition des résidus hydrophobes Leu6 et Phe13 du foldamère avec leur homologue. C : Mise en évidence de la superposition des résidus cycliques ACPC 3, 5, 7 avec des résidus analogues.

Ce dernier exemple, montre parfaitement que des unités cycliques peuvent être incluses dans les séquences de foldamères pour promouvoir une conformation particulière ou pour mimer une chaîne latérale en particulier.

### 3.1.2. Les foldamères aromatiques

Plus la chaîne aliphatique présente dans un monomère est longue, plus le nombre d'angles de torsion est potentiellement important. Cette augmentation du nombre de degrés de liberté rend moins prévisible le repliement de ces objets. L'utilisation de monomères cycliques permet de contraindre ces angles de torsion tout en conservant le nombre d'atomes de la chaîne principale (**Figure I.14.**).

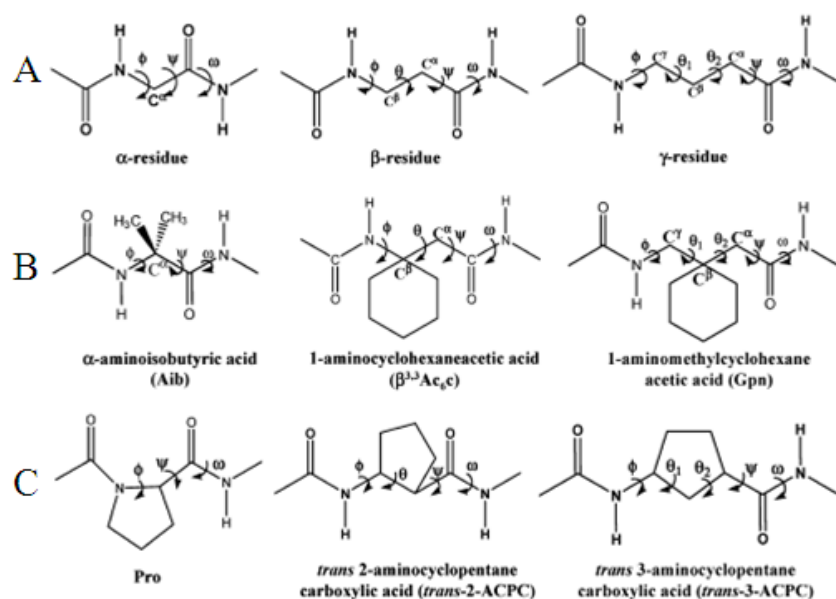


Figure I.14. : Résidus  $\alpha$ ,  $\beta$  et  $\gamma$  et exemples de résidus aromatiques correspondants.

A : Définition des angles de torsion pour les résidus  $\alpha$ ,  $\beta$  et  $\gamma$ . B : Résidus  $\alpha$ ,  $\beta$  et  $\gamma$  géminalement substitués.

C : Résidus  $\alpha$ ,  $\beta$  et  $\gamma$  possédant un cycle permettant de contraindre d'avantage encore les angles de torsion.

L'utilisation d'acides  $\gamma$ -aminés cycliques dans des séquences de foldamères aliphatiques permet d'obtenir des structures secondaires inédites. En effet l'équipe de Balaram (Vasudev *et al.* 2009) a pu obtenir des structures en coude de type I ou de type II (**Figure I.15.**) grâce à la succession d'un résidu  $\alpha$  et d'un résidu  $\gamma$  cyclique, la gabapentine (**Figure I.14.B Gpn**). Avec la mise en place de structures en coude, la préparation de séquences se repliant en feuillet  $\beta$  est plus aisée (**Figure I.15.**).

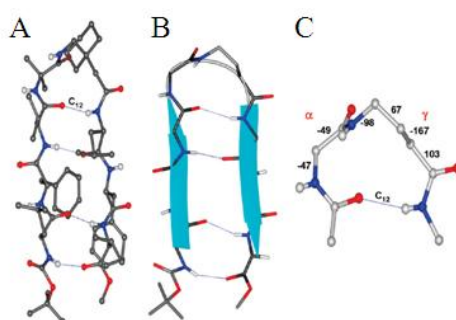


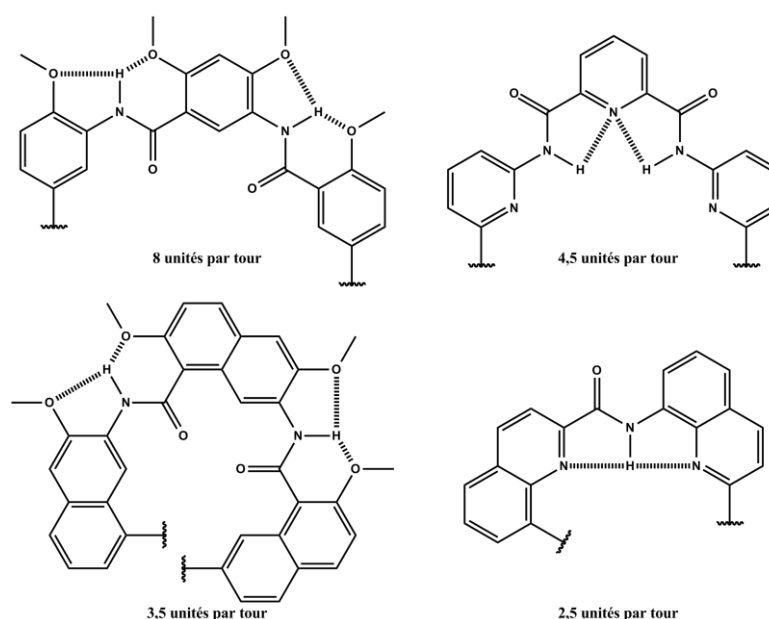
Figure I.15. : Structure en épingle à cheveux  $\beta$ .

A : Structure de l'épingle à cheveux  $\beta$  déterminée à partir de cristaux du peptide Boc-Leu-Phe-Val-Aib-Gpn-Leu-Phe-Val-OMe. B : Représentation de la même structure en ruban et bâtonnets sans les chaînes latérales. C : Détails du coude et de la liaison hydrogène C12 entre Gpn et Aib, avec les angles de torsion notés (Vasudev *et al.* 2009).



Grâce aux foldamères synthétiques et à la très grande variété de monomères (Zhang *et al.* 2012), les chimistes peuvent imaginer et concevoir de nouvelles structures.

Dans le cas de foldamères exclusivement aromatiques se repliant sous forme hélicoïdale, le diamètre de l'hélice et le nombre d'unités par tour sont déterminés par la courbure du monomère constitutif ainsi que par la nature du lien entre les différentes unités (amide ou urée). En fonction de la nature du cycle aromatique (benzène, pyridine, pyrimidine, quinoline...), de la position d'une unité par rapport à l'autre (*ortho*, *méta*, *para*), et des interactions non-covalentes mises en jeu, des structures secondaires différentes sont obtenues (**Figure I.16.**).



**Figure I.16. : Courbure de plusieurs oligomères aromatiques.**

Leur courbure dépend notamment du réseau de liaisons hydrogène représentées par des tirets, qui se développe à l'intérieur ou à l'extérieur de l'hélice et de l'unité aromatique utilisée (d'après Jiang *et al.* 2003).

C'est grâce à l'utilisation de plusieurs monomères de natures différentes que l'équipe d'Ivan Huc à Bordeaux a développé des séquences de foldamères présentant un rayon de courbure qui varie le long de la molécule, créant une cavité largement modulable et pouvant différencier, grâce entre autres à un réseau de liaisons hydrogène intermoléculaires, un L-tartrate d'un D-tartrate (excès diastéréomérique de plus de 99 %) en fonction de l'hélicité du foldamère (Ferrand *et al.* 2010) (**Figure I.17.**).

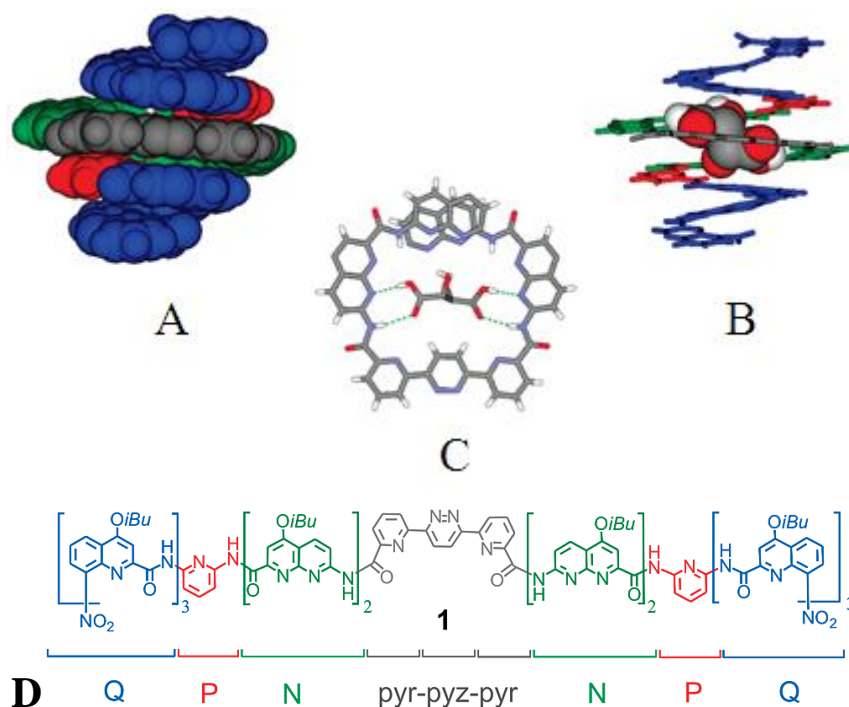


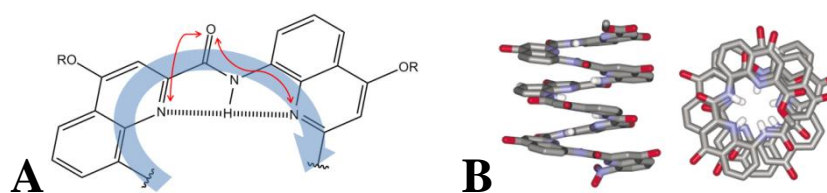
Figure I.17. : Foldamère hélicoïdal pouvant encapsuler une molécule d'acide tartrique L.

A et B : Représentations en sphères et en bâtonnets de la structure de l'acide tartrique L dans la cavité d'une hélice gauche (les chaînes isobutoxy (iBu) et les molécules d'eau ne sont pas représentées). C : Vue de haut du même complexe avec les liaisons hydrogène représentées en pointillés verts. D : La séquence du foldamère utilisant différents monomères quinolines (Q), pyridines (P) et 1,8-naphthyridines (N) permettant de créer la cavité, avec l'élément central composé des unités pyridine-pyridazine-pyridine.

### 3.2. Les oligoamides de quinolines

Les unités ou monomères quinolines sont des acides  $\delta$ -aminés, puisqu'il y a 4 atomes présents entre l'atome d'azote d'un résidu et le carbonyle du résidu suivant (**Figure I.18.** et **Figure I.19.**) (Jiang *et al.* 2003). Ces composés adoptent des conformations hélicoïdales repliées, exceptionnellement stables grâce à des liaisons hydrogène intramoléculaires parfaitement décrites et à des interactions d'empilement aromatique. Leur repliement est prévisible, leur synthèse aisée. Ils peuvent également être fonctionnalisés, ce qui fait de ces objets de bons candidats pour mimer les structures secondaires des macromolécules biologiques. Notons également que ces objets présentent une résistance accrue aux réactions d'hydrolyse ou de protéolyse. De récentes études ont également démontré la non-toxicité ainsi que la capacité à pénétrer les cellules humaines (HeLa, Huh-7 et Jurkat) que possèdent les foldamères (4-, 6-, et 8-mère) fonctionnalisés avec des charges positives (Iriondo *et al.* 2010).

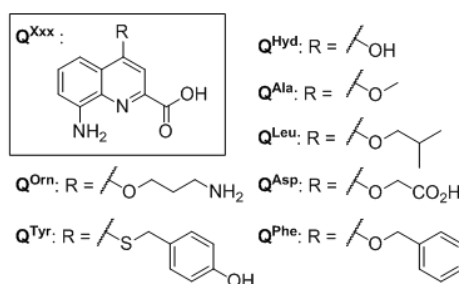
Le repliement hélicoïdal des oligoquinolines est induit par la conformation préférentielle « trans » de leurs fonctions amides (**Figure I.18.**). Celle-ci s'explique notamment par la répulsion entre le doublet de l'azote endocyclique et celui de l'oxygène du carbonyle. Ainsi, la courbure des oligoquinolines mène à des hélices qui présentent 2,5 unités par tour, un diamètre de 10 Å et un pas d'hélice d'environ 3,4 Å. De plus, la conformation hélicoïdale de ces foldamères est remarquablement stable dans l'eau et l'étude d'un octamère quinolines a montré que ce composé maintient sa conformation dans le DMSO à 120 °C (Jiang *et al.* 2003).



**Figure I.18. : Repliement d'un oligomère de quinolines**

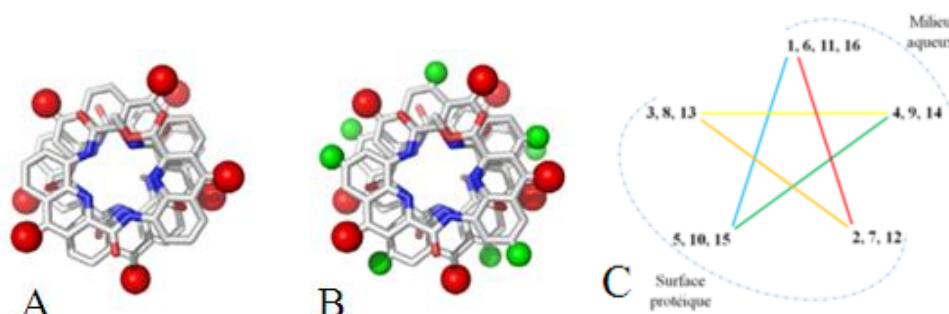
**A :** La courbure d'un oligomère de quinolines, due aux répulsions électrostatiques, matérialisées par les flèches rouges, et aux liaisons hydrogène intramoléculaires représentées par des tirets noirs. **B :** Structure cristallographique d'un octamère vu de côté et de dessus (les chaînes latérales ne sont pas représentées par souci de clarté).

Les monomères quinolines sont aisément fonctionnalisables. Les positions 4 et 6 de chaque sous-unité sont susceptibles de porter une fonction chimique différente. Les foldamères utilisés dans notre étude sont fonctionnalisés uniquement en position 4. Les groupements fonctionnels utilisés miment les chaînes latérales communément retrouvées dans les protéines. Nous noterons  $Q^{Xxx}$  (**Figure I.19.**) dans le reste de ce manuscrit tout monomère quinoline portant une chaîne latérale Xxx (Xxx faisant, tant que possible, référence à l'acide  $\alpha$ -aminé analogue).



**Figure I.19. : Monomères quinolines utilisés dans cette étude.**

Comme pour toute hélice biologique, certaines chaînes latérales seront plus particulièrement présentes pour participer à la solubilité de l'objet, alors que d'autres seront primordiales pour la reconnaissance de la cible au niveau moléculaire (**Figure I.20.**).



**Figure I. 20. : Distribution spatiale des positions fonctionnalisables.**

**A et B : Structure cristallographique d'un octamère de quinolines. La position 4 de chaque monomère est représentée par une sphère rouge ; la position 6 est matérialisée par une sphère verte. C : Distribution de la position 4 en étoile à 5 branches. Les nombres correspondent à la position de l'unité quinoline dans le foldamère.**

Ces oligoquinolines sont des peptides orientés, c'est-à-dire qu'ils possèdent une extrémité N-terminale et une extrémité C-terminale distinctes.

Ces composés peuvent être synthétisés en solution, à la manière d'une synthèse peptidique classique, par une approche convergente avec de bons rendements, en utilisant une activation par des agents de couplage ou par un intermédiaire chlorure d'acyle (Jiang *et al.* 2003). Cette méthode de synthèse était, lors du début des travaux exposés dans ce manuscrit, la méthode de choix car il s'agissait alors de la méthode la plus efficace et la plus économique, en terme de temps, pour accéder à de longs foldamères de première génération composés de bloc de la même unité (même chaîne latérale). Depuis octobre 2012, la synthèse sur support solide des oligoquinolines (Baptiste *et al.* 2010) fut développée et finalement généralisée. Elle est à la base des nombreux foldamères de seconde génération synthétisés et utilisés dans notre travail. Cette technique s'est révélée être plus rapide, plus économique et plus efficace.

Le passage d'une technique de synthèse à l'autre, a toutefois nécessité des ajustements importants dans la façon de concevoir les foldamères. En effet la synthèse en solution s'effectue de l'extrémité N-terminale vers l'extrémité C-terminale contrairement à la synthèse sur support solide qui s'effectue en sens inverse.

Les foldamères, qu'ils soient aliphatiques, aromatiques, ou hybrides, sont des molécules innovantes capables d'adopter une conformation spatiale prévisible. La taille moyenne, la conformation stable, la facilité avec laquelle il est possible de les synthétiser et de les fonctionnaliser, la résistance accrue aux protéases et leur non-toxicité font de ces molécules d'excellents candidats pour inhiber des IPP.

## Conclusion

Les interactions protéine-protéine sont au cœur de tous les processus biologiques et chacune d'entre elles est unique (taille, relief, polarité, ...). Comme nous l'avons montré, les IPP sont des cibles thérapeutiques présentant un grand potentiel lorsqu'on les compare aux sites actifs des enzymes. Leurs analyses montrent cependant que leur inhibition n'est pas aussi aisée que pourrait l'être celle d'une enzyme. Les méthodes d'analyse et de prédiction à tous les niveaux de résolution (atomique, énergétique, cellulaire) apportent des informations indispensables à la conception d'inhibiteur de haute affinité et de forte spécificité.

La communauté scientifique a très vite compris que le paramètre indispensable à l'inhibition d'une IPP était la complémentarité de l'inhibiteur pour la surface à reconnaître. Aussi a-t-elle utilisé une variété de molécules support structuralement définies et exposant des chaînes latérales variables. Les anticorps et les protéines miniatures en sont les parfaits exemples et ont démontré leur capacité à inhiber des IPP.

Les chimistes se sont lancés dans la création d'une nouvelle classe de molécules possédant les qualités des protéines (taille, solubilité, non-toxicité, ...) tout en supprimant leurs défauts (sensibilité aux protéases, difficulté à prédire la structure, ...) : les foldamères. Ces molécules sont ainsi toutes désignées comme des candidats potentiels pour l'inhibition des IPP.

Nous nous sommes particulièrement intéressés aux foldamères de quinolines, qui adoptent une structure hélicoïdale parfaitement définie et prévisible et qui peut être fonctionnalisée par des chaînes latérales très variables. Au cours de ces travaux nous nous sommes appliqués à démontrer la capacité des foldamères à reconnaître une surface protéique spécifique en utilisant la cristallographie des rayons X pour optimiser le choix de ces chaînes latérales.

## Chapitre II

### Inhibition du complexe IL-4R $\alpha$ /IL-4



## Introduction

Le travail réalisé au cours de cette thèse a pour but de démontrer qu'un foldamère de type oligoamide de quinoline correctement fonctionnalisé, est capable d'interagir avec une surface protéique donnée. Cette étape, nous l'avons souligné, est en effet la première à valider avant de démontrer que de tels composés peuvent inhiber les interactions protéine-protéine. Deux protéines (l'interleukine 4 humaine et l'anhydrase carbonique humaine II) ont été retenues pour illustrer cette possible reconnaissance. Ce premier chapitre consacré aux résultats obtenus lors de ce mémoire fait état des travaux menés sur l'interleukine 4 (IL-4). Le but recherché est de montrer qu'un foldamère est susceptible de reconnaître la surface de l'IL-4 impliquée dans l'interaction avec son récepteur de haute affinité, IL-4R $\alpha$ . L'un des partenaires académiques du programme Foldappi, dans le cadre duquel a été conduit ce travail, est l'équipe des professeurs Walter Sebald et Thomas Müller de l'Université de Würzburg (Allemagne). Le travail conduit sur l'IL-4 a fortement bénéficié de leur expertise puisqu'ils sont à l'origine de nombreuses études tant biochimiques, structurales que fonctionnelles sur cette protéine et ses différents partenaires. Après une brève présentation des données bibliographiques concernant le système IL-4R $\alpha$ /IL-4, nous exposons dans ce chapitre la démarche suivie et les résultats obtenus.

## 1. L'IL-4 : Fonctions, Structures et inhibition

---

### 1.1. Physiologie de l'IL-4

#### *1.1.1. Rôles et fonctions de l'IL-4*

L'interleukine 4 est une macromolécule biologique pléiotropique découverte dans les années 1980 et fait partie de la famille des cytokines de type I (Luzina *et al.* 2012). Ces protéines ou glycoprotéines, permettant la communication intercellulaire, sont synthétisées par les cellules du système immunitaire et dont le rôle majeur est la régulation de la réponse immunitaire dans les allergies (Guenounou 1998), autres rhinites et conjonctivites, ainsi que



dans les inflammations plus sérieuses comme l'asthme (Oh *et al.* 2010), le cancer (Hallett *et al.* 2012) ou le SIDA (Vasilescu *et al.* 2003).

Le type cellulaire produisant initialement l'IL-4 n'a pas encore été clairement défini, mais une étude relativement récente (Sokol *et al.* 2008) suggère qu'il pourrait s'agir des lymphocytes basophiles. Dans ce cas, l'IL-4 régule positivement la différenciation des lymphocytes T (LT) auxiliaires naïfs (Th0) en cellules Th2 activées. Ce processus d'activation ne se produit que lorsqu'un antigène (allergène par exemple) est absorbé par une cellule présentatrice d'antigène (APC), et que cette dernière présente cet antigène à la cellule Th0. L'activation induit la production de l'IL-4 mais aussi d'autres cytokines comme l'IL-5, l'IL-10 et l'IL-13 (Seder *et al.* 1992). L'IL-4 est également produite par les lymphocytes mastocytaires, basophiles, et éosinophiles (Nelms *et al.* 1999). Cette surabondance subite de cytokines (en particulier d'IL-4) provoque la croissance, la différenciation et le maintien des lymphocytes T et B (Hu-Li *et al.* 1987). La différenciation des lymphocytes B (LB) en plasmocytes inclut notamment l'amplification de l'expression du complexe majeur d'histocompatibilité (CMH) de type II (Nelms *et al.* 1999), la régulation positive de l'expression du récepteur de l'IL-4 (Ohara et Paul 1988) ainsi que l'induction de la commutation isotypique des immunoglobulines vers les anticorps IgE dont la production est accélérée (Ohara et Paul 1988). Jouant un rôle important dans les réactions d'hypersensibilité immédiate, ces anticorps sont essentiels à la réponse inflammatoire allergique.

L'IL-4 n'intervient pas seulement au niveau des LT et des LB. En effet, elle active différentes voies de signalisation :

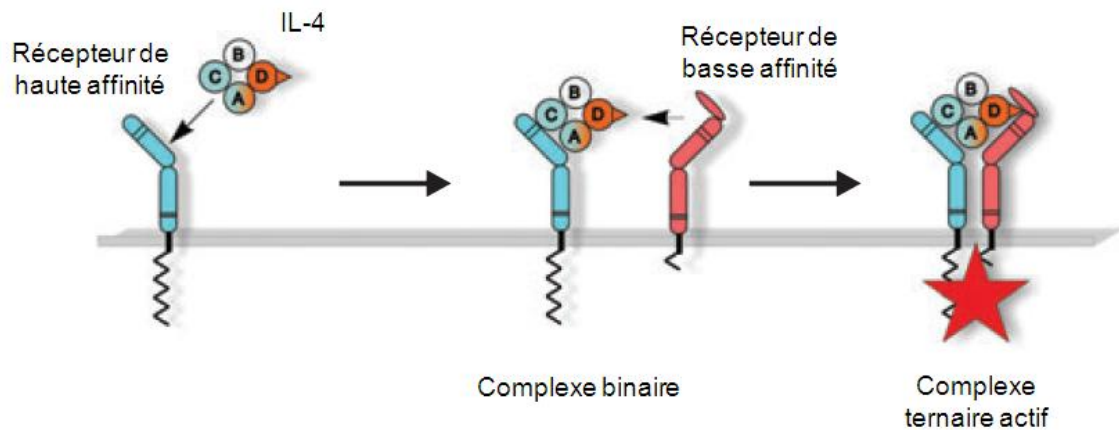
- Elle active de manière alternative les macrophages tissulaires en cellules M2 plutôt qu'en cellules M1. L'activation de ces cellules participe à la réduction de l'inflammation pathologique ainsi qu'à la cicatrisation (Aster *et al.* 2009).
- Elle agit sur les cellules musculaires lisses et sur les cellules épithéliales (Luzina *et al.* 2012). Elle provoque en particulier l'augmentation de la production de mucus par les cellules épithéliales respiratoires caliciformes.

### 1.1.2. Mécanismes d'action de l'IL-4

Comprendre et inhiber les mécanismes par lesquels l'IL-4 agit nécessite de s'intéresser aux récepteurs de cette cytokine.

Pour activer les voies de signalisation qui lui sont propres, l'IL-4 doit être reconnue par son récepteur, présent à la surface de très nombreux types cellulaires. L'IL-4 va provoquer

l'hétérodimérisation de son récepteur de haute affinité (150 pM) IL-4R $\alpha$  avec la chaîne commune  $\gamma_c$ , récepteur de basse affinité (4  $\mu$ M) (Zhang *et al.* 2002) (**Figure II.1.**). Des études ont montré que  $\gamma_c$  reconnaît le complexe IL-4R $\alpha$ /IL-4 (Letzelter *et al.* 1998). Bien que l'association du récepteur  $\gamma_c$  au complexe en formation n'augmente que très peu l'affinité de l'IL-4 pour son récepteur de haute affinité, elle reste nécessaire à l'activation de la voie de signalisation (Russell *et al.* 1993).



**Figure II.1. : Séquence d'activation du complexe ternaire. (D'après Sebald *et al.* 2010 et Zhang *et al.* 2002)**

Certaines cellules n'expriment pas  $\gamma_c$  mais restent toutefois sensibles à l'IL-4 (Dawson *et al.* 1997), ce qui indiquerait qu'une autre chaîne peut être utilisée en lieu et place du récepteur de basse affinité  $\gamma_c$ . De récentes études indiquent que la chaîne IL-13R $\alpha$ 1 est ce récepteur, lorsque  $\gamma_c$  n'est pas exprimée, en particulier dans les cellules non hématopoïétiques.

Il est important de noter que l'IL-4 est très proche structuralement et fonctionnellement de l'IL-13. En effet l'IL-13, pour activer ses voies de signalisation, doit également former un complexe ternaire composé des récepteurs IL-4R $\alpha$  et IL-13R $\alpha$ 1 (LaPorte *et al.* 2008). Le recoupement des systèmes de récepteur va dans le sens de l'existence d'une synergie fonctionnelle des deux interleukines. La structure d'IL-13 a été déterminée par RMN (Moy *et al.* 2001), et celle de complexe ternaire IL4-R $\alpha$ /IL-13/IL-13R $\alpha$ 1 par diffraction des rayons X (Laporte *et al.* 2008).

Ainsi l'IL-4 peut activer les processus cellulaires grâce à deux combinaisons de récepteurs, celui de type I qui inclut IL4-R $\alpha$ /IL-4/ $\gamma_c$  et celui de type II qui inclut IL4-R $\alpha$ /IL-4/IL-13R $\alpha$ 1.

## 1.2. Structures de l'IL-4 et du complexe IL-4R $\alpha$ /IL-4

### 1.2.1. Structure de l'IL-4

L'IL-4 est une protéine globulaire d'environ 15 kDa dont la séquence primaire comprend 129 acides aminés, lorsque le peptide signal est clivé. Cette protéine est essentiellement composée d'un fagot de quatre hélices  $\alpha$  auquel est accolé un feuillet  $\beta$  constitué de 2 brins anti-parallèles (**Figure II.2.**) Trois ponts disulfures stabilisent la structure : C3-C127, C24-C65 et C46-C99 (Vaz *et al.* 2005). La structure de cette protéine a d'abord été déterminée par RMN (Code PDB : 1ITL) (Smith *et al.* 1992) puis par diffraction des rayons X (Codes PDB : 1RCB et 2B8U) (Wlodawer *et al.* 1992 et Kraich *et al.* 2006).

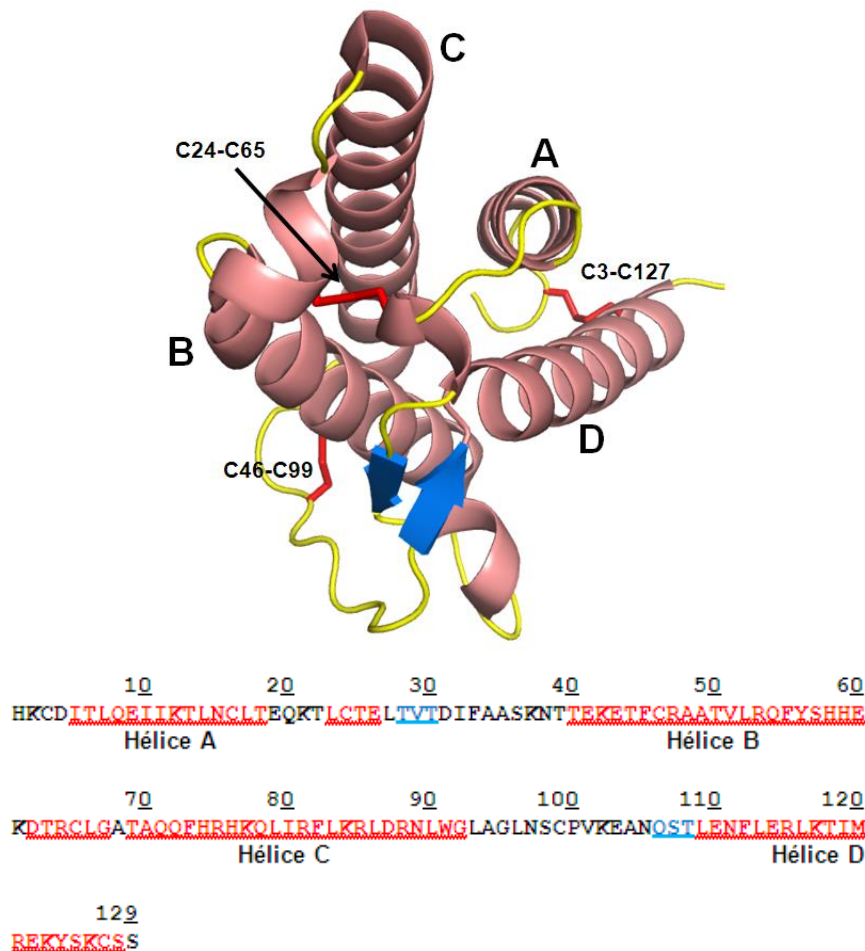
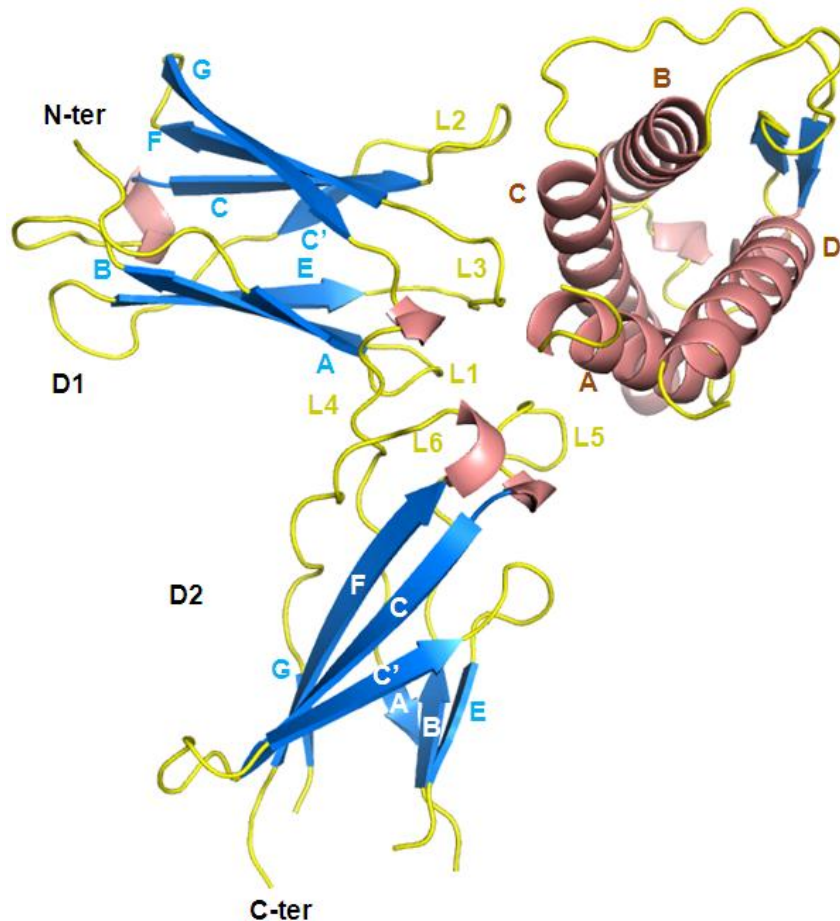


Figure II.2. : Structure de l'IL-4 (Code PDB : 2B8U)

Les trois ponts disulfures de la protéine sont représentés en bâtonnets rouge et la position des cystéines est indiquée. La séquence est reportée avec le même code couleur (rouge pour les hélices  $\alpha$ , et bleu pour les feuillets  $\beta$ ).

### 1.2.2. Le complexe IL-4R $\alpha$ /IL-4

A notre connaissance, la structure du récepteur de haute affinité de l'IL-4 seul n'a jamais été déterminée. Au contraire, la structure du complexe binaire IL-4R $\alpha$ /IL-4 (Hage *et al.* 1999) et celles des complexes ternaires IL-4R $\alpha$ /IL-4/ $\gamma_c$  et IL-4R $\alpha$ /IL-4/IL-13R $\alpha$ 1 (LaPorte *et al.* 2008) ont été déterminées par diffraction des rayons X. Les coordonnées atomiques ont été déposées à la PDB. Les codes d'entrée sont respectivement les suivants : 1IAR, 3BPL et 3BPN. Nous décrivons ici la structure du complexe IL-4/IL-4R $\alpha$ , déterminée à la résolution de 2,30 Å, qui permet, à elle seule, de préciser la surface d'interaction de l'IL-4 avec son récepteur.



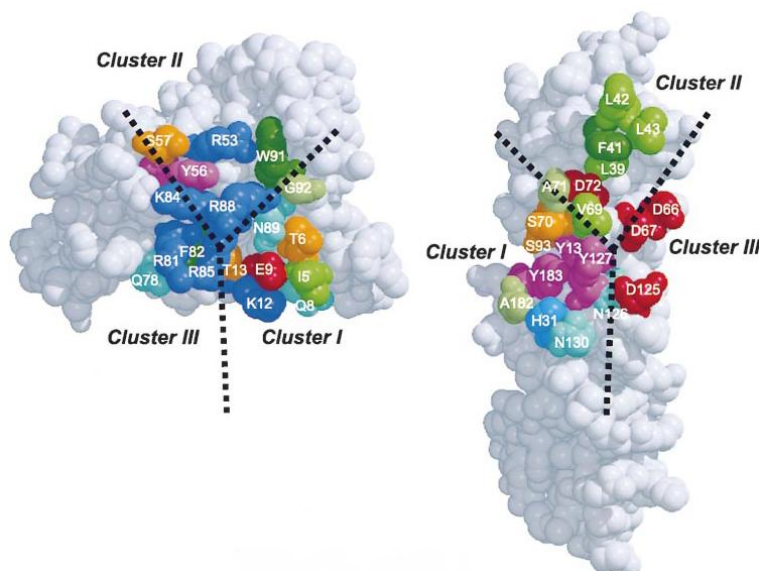
**Figure II.3. : Structure du complexe binaire IL-4R $\alpha$ /IL-4 (Code PDB : 1IAR).**

IL-4R $\alpha$  est composé de deux domaines D1 (résidus 1 à 91) et D2 (résidus 97 à 197). Ces domaines, topologiquement proches des modules de fibronectine de type III, adoptent une conformation en sandwich de deux feuillets  $\beta$  antiparallèles arrangés en trois (A, B, E) et

quatre brins (C, C', F et G) (**Figure II.3**). Le domaine D1 est le seul à posséder des cystéines (6 au total) engagées dans trois ponts disulfures (C9-C19, C49-C61 et C29-C59). La reconnaissance de l'IL-4 par IL-4R $\alpha$  s'effectue essentiellement par 6 boucles (L1 à L6) incluant respectivement les résidus 11-16, 39-47, 64-74, 90-102, 119-133 et 179-194. L2 interagit avec l'hélice B de l'IL-4 ; L1 et L3 avec l'hélice C et L4, L5 et L6 avec l'hélice A.

Le repliement de l'IL-4 adopté dans le complexe IL-4R $\alpha$ /IL-4 est relativement proche de celui observé lorsque la protéine est cristallisée seule. Le r.m.s.d calculé sur la position des 129 C $\alpha$  est de 1,1 Å. Des écarts supérieurs à 2 Å sont observés. Néanmoins, ceux-ci concernent essentiellement soit les 5 premiers C $\alpha$  de la partie N-ter, soit des boucles (résidus : 35-37, 66-67) localisées loin des zones de contact avec le récepteur.

L'interface existante entre l'IL-4 et son récepteur couvre une surface légèrement supérieure à 1 400 Å<sup>2</sup> (LaPorte *et al.* 2008). Sur les 18 résidus de l'IL-4 présents à l'interface, 6 enfouissent 80% de leur surface lors de la formation du complexe. Ces résidus sont principalement localisés sur les hélices A et C de l'IL-4 et sont majoritairement des résidus polaires ou chargés. L'épitope complémentaire sur le récepteur concerne principalement les boucles L2, L3 et L5. Il est caractérisé par un patch hydrophobe, un autre principalement constitué d'aspartates et un dernier formé d'acides aminés polaires non chargés (Tyr, Ser et Asn).

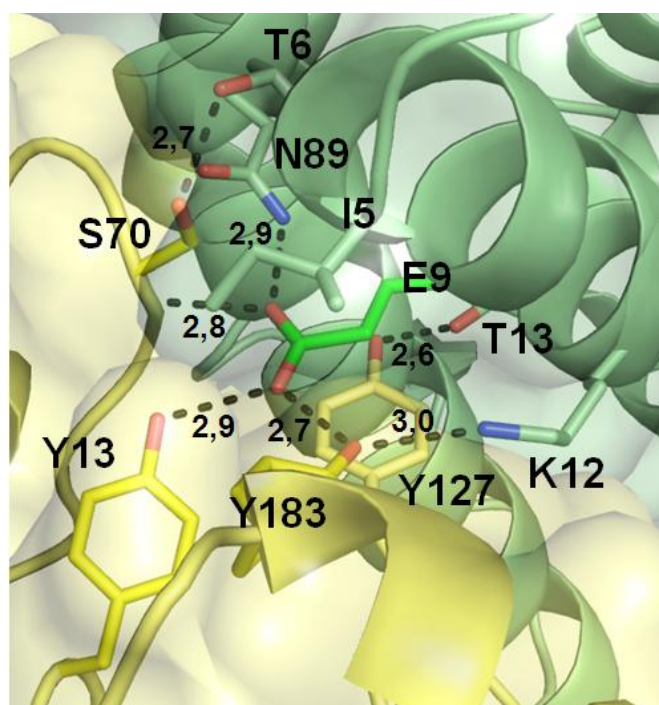


**Figure II.4. : Vue du complexe IL-4R $\alpha$ /IL-4 ouvert.**

L'épitope de l'IL-4 à gauche et celui du récepteur à droite sont divisés en trois groupes ou *clusters* distincts (Hage *et al.* 1999).

Les épitopes présents sur l'IL-4 et son récepteur sont chacun composé de trois clusters distincts qui interagissent par paire. Deux d'entre eux présentent un caractère amphipathique avec une couronne externe hydrophobe et un cœur polaire (**Figure II.4.**).

Le premier *cluster* est centré sur E9 de l'IL-4 dont le groupement carboxylate établit trois liaisons hydrogène avec les résidus Y13, Y183 et S70 du récepteur. La chaîne latérale de E9 semble être maintenue dans une orientation favorable grâce à une liaison hydrogène avec la chaîne latérale du résidu N89 de l'IL-4. D'autres liaisons hydrogène (entre les résidus K12, T13, N89 et T6 de l'IL-4 et les résidus précédemment cités du récepteur) participent également à la complémentarité des épitopes. E9 établit en outre des contacts de van der Waals avec la chaîne latérale Y127 du récepteur, sur laquelle il est littéralement couché. Les résidus hydrophobes Y13, Y183, du récepteur et I5 de l'IL-4 constituent la couronne hydrophobe, et accroissent la force des interactions autour de E9 (**Figure II.5.**).



**Figure II.5. : Réseau de liaisons hydrogène du cluster I centré autour du glutamate E9 de l'IL-4.**  
L'IL-4 est en vert et le récepteur IL-4R $\alpha$  est en jaune. Les distances sont rapportées en angströms.



L'IL-4 est en vert et le récepteur IL-4R $\alpha$  est en jaune. Les distances sont en angströms.

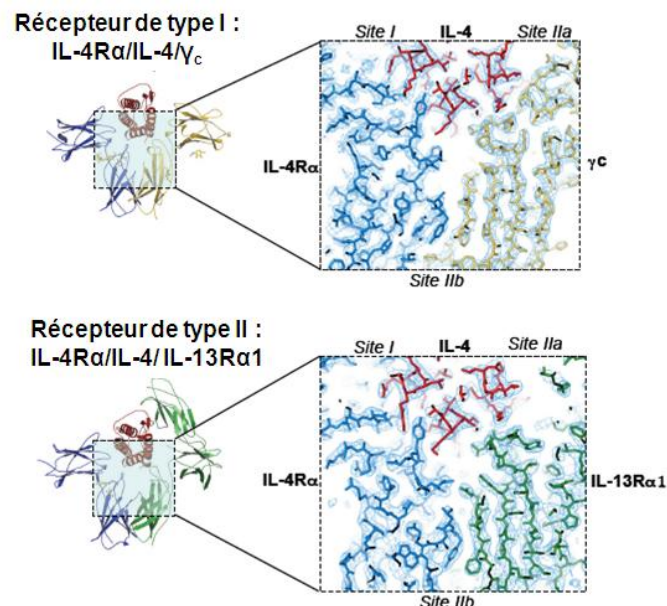


L'IL-4 est en vert et le récepteur IL-4R $\alpha$  est en jaune. Les distances sont rapportées en angströms.

Des études de mutagenèse dirigée ont montré (Hage *et al.* 1999) que les clusters I et II sont majoritairement responsables de l'énergie d'interaction entre les deux partenaires. La mutation des résidus E9 et R88 provoque la chute de la constante d'affinité de deux à trois ordres de grandeur alors que des mutations des résidus périphériques tels que N89, I5 ou T13 ne diminuent la constante d'affinité que d'un peu moins d'un ordre de grandeur.

### 1.2.3. Les complexes ternaires de type I et de type II

Le but recherché à ce stade de notre travail était de préciser les surfaces d'interaction de l'IL-4 avec son récepteur de haute affinité, ce que nous avons fait dans le paragraphe précédent. Il est toutefois intéressant de vérifier si cette surface n'est pas fortement modifiée lorsque le troisième partenaire,  $\gamma_c$  ou IL-13R $\alpha$ 1, interagit avec le complexe binaire IL-4R $\alpha$ /IL-4. La structure des complexes ternaires (codes PDB : 3BPL et 3BPN), déterminée par diffraction des rayons X à la résolution de 2,95 Å et 3,0 Å respectivement (**Figure II.8.**) (Laporte *et al.* 2008), permet de répondre à cette interrogation. Le récepteur  $\gamma_c$  possède une structure comparable à celle du récepteur IL-4R $\alpha$  : ces deux récepteurs sont, l'un comme l'autre, composés de deux domaines fibronectine de type III. Le récepteur IL-13R $\alpha$ 1 quant à lui possède trois domaines fibronectine de type III.



**Figure II.8. :** Définition des différents sites d'interactions entre partenaires suivant le type de récepteur L'IL-4 est en rouge, IL-4R $\alpha$  est en bleu, la chaîne  $\gamma_c$  est en jaune et IL-13R $\alpha$ 1 est en vert (D'après Laporte *et al.* 2008).



Le site I est défini comme étant l'interface entre l'IL-4 et son récepteur de haute affinité. Les sites IIa et IIb sont les interfaces entre l'IL-4 et son récepteur de basse affinité ( $\gamma_c$  ou IL-13R $\alpha$ 1) et entre les deux récepteurs de haute et de basse affinité. Les principales caractéristiques du site I suivant le type de récepteur, sont mentionnées dans le **Tableau II.1**.

	Complexe binaire	Complexe ternaire de type I	Complexe ternaire de type II
Surface enfouie ( $\text{\AA}^2$ )	1 550	1 570	1 390
Nombre de contacts hydrophobes ( $< 5 \text{\AA}$ )	4	4	3
Nombre de contacts de van der Waals ( $< 4,5 \text{\AA}$ )	25	25	23
Nombre de liaisons hydrogène ( $< 3,5 \text{\AA}$ )	18	22	15
Nombre de contacts ioniques ( $< 6 \text{\AA}$ )	8	7	6

**Tableau II.1. : Principales caractéristiques du site I en fonction du type de récepteur impliqué.**

Les valeurs mentionnées sont issues de calculs réalisés avec les serveurs PISA (Krissinel et Henrick 2007) et PIC (Tina *et al.* 2007).

En utilisant comme critères d'interaction des distances inférieures ou égales à  $5 \text{\AA}$  pour des contacts hydrophobes, et des distances inférieures ou égales à  $3,5 \text{\AA}$  pour des liaisons de type hydrogène ou polaire, les résidus de l'IL-4 impliqués dans l'interaction avec IL-4R $\alpha$  ne varient guère en présence ou non de  $\gamma_c$  ou de IL-13R $\alpha$ 1. Les modifications, mineures au demeurant, ne remettent donc pas en cause, lorsqu'est formé le complexe ternaire, la surface de reconnaissance de l'IL-4 vis-à-vis de son récepteur de haute affinité IL-4R $\alpha$  définie à partir de la seule structure du complexe binaire.

### 1.3. Stratégies d'inhibition du complexe

Pour bloquer tous les mécanismes cellulaires déclenchés par la présence de l'IL-4 et/ou de l'IL-13, l'inhibition de complexes intra-cellulaires propres à une voie de signalisation activée par une de ces deux cytokines semble être un projet stimulant mais vallonné de nombreux défis. L'inhibition de la formation du complexe ternaire au niveau de la chaîne  $\gamma_c$  est problématique puisque cette chaîne est également utilisée par les interleukines 2, 7, 9, 15 et 21 (Meazza *et al.* 2011). Cela induirait de nombreux effets secondaires. L'inhibition du

complexe au niveau du récepteur IL-13R $\alpha$  semblerait thérapeutiquement plus efficace, mais n'agirait pas sur la différenciation des lymphocytes Th0 en Th2 par exemple puisque les cellules hématopoïétiques possèdent le second système de récepteurs impliquant  $\gamma_c$ . Ainsi l'inhibition du complexe que forme l'IL-4/13 avec IL-4R $\alpha$  serait la plus pertinente. Les IL-4 et IL-13 sembleraient être les seules cytokines connues à se lier au récepteur IL-4R $\alpha$  (Mueller *et al.* 2002) ce qui éviterait l'apparition d'effets secondaires. A ce jour, des anticorps, des protéines mutantes et des protéines miniatures ont été développés comme antagonistes de l'IL-4 ou de l'un de ses récepteurs.

### 1.3.1. Par des anticorps

Des anticorps monoclonaux dirigés contre l'IL-4 (Tony *et al.* 1994) ou contre le récepteur IL-4R $\alpha$  (Reusch *et al.* 1994 et Kakkar *et al.* 2011) ont été développés et semblent efficaces *in vitro* et *in vivo*. Notons que ces deux types d'anticorps ne sont pas fonctionnellement équivalents. Seul celui dirigé contre le récepteur IL-4R $\alpha$  inhibe la formation des complexes avec l'IL-4 et l'IL-13. De manière surprenante les anticorps anti IL-4 peuvent accroître ses effets biologiques (Finkelman *et al.* 1993). En effet l'association de l'anticorps avec l'interleukine augmente son temps de demi-vie dans le sérum en la protégeant de la dégradation protéolytique et de son élimination rénale, potentialisant son effet lorsqu'elle est libre.

### 1.3.2. Par la mutation de protéines

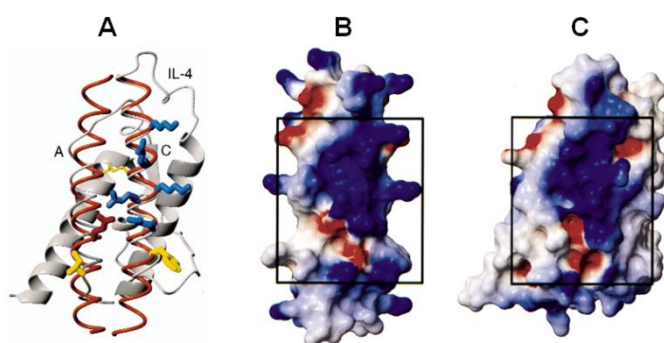
Une alternative aux anticorps, grosses molécules biologiques relativement chères à produire, serait d'utiliser une version soluble du récepteur qui piègerait l'IL-4 produite et l'empêcherait de se fixer sur le récepteur membranaire pour activer les voies de signalisation (Renz *et al.* 1996). Les domaines extracellulaires des récepteurs de cytokine peuvent être produits biologiquement par épissage alternatif des ARNm (Heaney et Golde 1996). L'avantage de cette méthode est de faire appel à des molécules physiologiques qui ne déclencheront pas de réponse immunitaire. Il est cependant nécessaire de noter qu'une telle démarche est susceptible de potentialiser les effets de l'IL-4 comme le font les anticorps anti IL-4. Un variant soluble du récepteur IL-4R $\alpha$  a prouvé son efficacité chez l'homme (Borish 2001).

Une autre alternative serait de muter certains résidus de l'IL-4 pour obtenir une protéine antagoniste. Ainsi la mutation des résidus R121 et Y124 en aspartate, tout en

conservant une forte affinité de l'IL-4 mutante pour son récepteur de haute affinité ( $\sim 100$  pM), ne lui permet plus de se complexer avec le récepteur de basse affinité quel qu'il soit (Tony et al. 1994). Ainsi la formation du complexe ternaire est bloquée et la voie de signalisation est inhibée. Ce mutant fait l'objet d'un essai clinique de phase III qui a été validé comme traitement contre l'asthme allergique (Wenzel *et al.* 2007).

### 1.3.3. Par des protéines miniatures

Une démarche différente a été menée par l'équipe de Luis Serrano (Domingues *et al.* 1999). Elle consiste à utiliser le domaine « leucine zipper » ( $\sim 3,5$  kDa) du facteur de transcription GCN4 (code PDB : 1DGC ; König et Richmond 1993) comme domaine sur lequel pourrait être reproduit l'épitope de l'IL-4 impliqué dans les interactions IL-4/IL-4R $\alpha$ . La structure de GCN4 a été choisie car elle comprend deux hélices organisées en superhélice (domaine *coiled-coil*) susceptibles de mimer les hélices A et C de l'interleukine 4. Après plusieurs essais, un mime a été obtenu. Il inclut 8 résidus de l'IL-4 (I5 et E9 de l'hélice A et K77, R81, K84, R85, R88, W91 de l'hélice C) et présente une affinité pour le récepteur de  $26 \mu\text{M}$ . L'ajout d'un pont disulfure intra-moléculaire (**Figure II.9.**) augmentant la stabilité de la molécule, permet d'accroître l'affinité jusqu'à une valeur de  $5 \mu\text{M}$ . Cette molécule ne présentant malgré tout qu'une affinité moyenne pour entrer en compétition avec le ligand naturel du récepteur, les auteurs ont suggéré d'introduire de nouveaux résidus (T13, Q89 et R53) afin d'améliorer encore l'affinité (Domingues *et al.* 1999).



**Figure II.9. : Comparaison structurale entre l'IL-4 et son mime.**

**A :** Superposition des structures de l'IL-4 en gris et du mime en rouge. Les chaînes latérales représentées sont celles qui ont été transférées sur GCN4, le pont disulfure est représenté en fin bâtonnets jaune. **B et C :** Comparaison de la distribution des charges à la surface du mime (B) et de l'IL-4 (C) responsable de l'interaction avec le récepteur.

Ainsi l'approche s'appuyant sur des anticorps et sur des protéines mutées ou miniatures a l'avantage d'utiliser des objets très affins. Les inconvénients majeurs de ces stratégies sont la potentialisation de l'effet de l'IL-4 et la taille relativement importante des anticorps, inconvénients que n'ont pas les mimes de l'IL-4. L'utilisation de mimes est de ce point de vue beaucoup plus pertinente. L'exemple donné plus haut concernant GCN4 semble être très prometteur, mais nécessite encore des étapes d'optimisation. Aussi nous sommes-nous intéressés à l'utilisation de foldamères d'oligoamides de quinolines pour reconnaître la surface de l'IL-4 et finalement inhiber la formation du complexe IL-4R $\alpha$ /IL-4.

## 2. Inhibition du complexe IL-4/IL-4R $\alpha$ par des foldamères

### 2.1. Stratégie générale

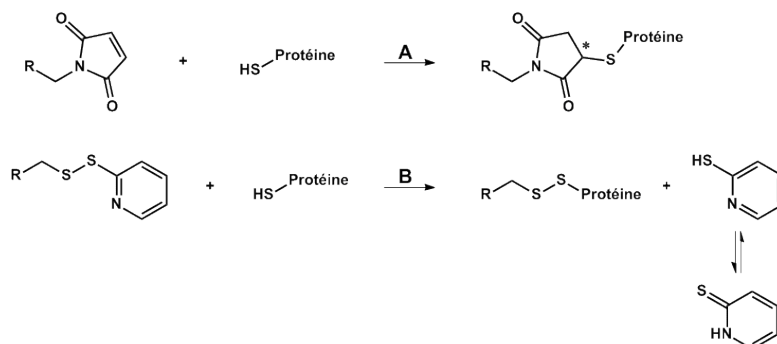
La stratégie choisie pour inhiber le complexe que forme l'IL-4 avec son récepteur de haute affinité IL-4R $\alpha$  consiste à concevoir un foldamère portant des chaînes latérales complémentaires de la surface de l'IL-4 impliquée dans la reconnaissance de son récepteur. Ainsi à terme, le foldamère pourra être utilisé pour piéger l'IL-4 et l'empêcher d'interagir avec son récepteur.

Plutôt que d'opter pour une approche par criblage d'un grand nombre de séquences de foldamère, le projet s'appuie sur une approche rationnelle qui préfère concevoir une molécule cible par modélisation moléculaire, puis valider et optimiser les interactions protéine/foldamère par cristallographie. L'idée d'approcher, par un système d'ancrage, le foldamère de la surface cible à reconnaître permettra, nous l'espérons, de favoriser l'établissement des interactions recherchées. La stratégie choisie pour promouvoir les contacts entre les deux molécules consiste à introduire par mutagenèse dirigée une cystéine dans la séquence de l'IL-4. Cette cystéine servira de point d'ancrage du foldamère via une liaison covalente. Cette stratégie d'ancrage ou *tethering* a déjà prouvé son utilité (Erlanson *et al.* 2004). Pour montrer que l'interaction attendue se produit, le foldamère devra posséder un certain degré de liberté de mouvement. Ce dernier est conféré par un espaceur, intercalé entre le groupement thiol de la cystéine et le foldamère.

## 2.2. Mutation, liaison chimique et espaceur

La position de la mutation qui permettra d'ancrer le foldamère est choisie par modélisation moléculaire. Cette position doit exposer le foldamère au solvant et être proche de la surface à reconnaître. Le résidu S16 a été considéré comme pertinent d'une part car sa localisation répond aux critères précédents, et d'autre part car la mutation d'une sérine en cystéine ne devrait pas avoir d'effets importants sur la stabilité de la protéine.

Plusieurs fonctions chimiques peuvent réagir avec une fonction thiol, telle par exemple une seconde fonction thiol activée, ou une fonction maléimide. La fonction thiol formant un pont disulfure avec la protéine semble, dans notre cas, plus adaptée dans le sens où elle n'implique pas, contrairement au couplage d'une fonction thiol à un maléimide, de carbone asymétrique (**Figure II.10.A et B**). L'existence d'un tel carbone induirait en effet l'existence de deux énantiomères qu'il faudrait séparer pour mener à bien la cristallogénèse des adduits. Nous avons cependant utilisé les deux types de liaison dans la suite de ce travail dans le cas éventuel où l'une d'entre elle ne s'avèrerait pas efficace.



**Figure II.10. : Schémas réactionnels d'un maléimide et d'un thiol activé avec une fonction thiol**

**A :** Le thiol réagit sur la double liaison de la fonction maléimide pour donner un thioéther (réaction de Michael) comportant un carbone asymétrique (\*). Au cours de la réaction d'échange. **B,** un anion thiolate (sur la protéine dans notre cas) déplace un soufre du pont disulfure. Le groupement partant 2-mercaptopyridine est stabilisé par isomérisation.

Les espaceurs utilisés sont de type alkyle et comportent une liaison amide pour des besoins de synthèse. Deux molécules **A** et **B** (**Figure II.11.**), synthétisées par l'équipe d'Ivan Huc à Bordeaux, ont été utilisées dans la suite de ce travail. Elles présentent le même espaceur lié au foldamère par son extrémité C terminale, et se différencient uniquement par les fonctions portées par le foldamère (tétramère). Le choix des fonctions portées par chaque

tétramère assure une bonne solubilité des molécules dans l'eau. En aucun cas, ce choix ne résulte d'une analyse de la surface protéique à reconnaître.

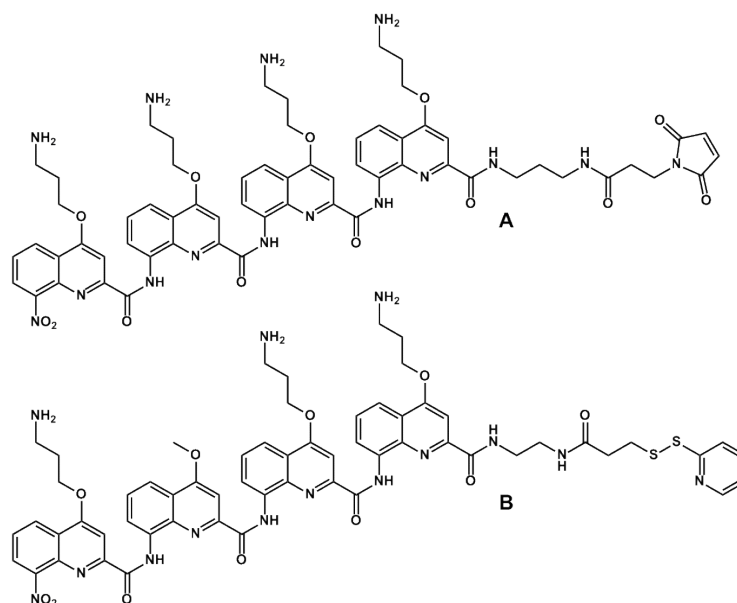


Figure II.11. : Molécules A et B portant soit un maléimide soit une fonction thiol activée.

### 3. L'adduit foldamère/IL-4 S16C

Avant toute tentative d'accrochage d'un foldamère à la protéine, nous avons dû produire la protéine mutée. Les étapes de production et de purification sont issues des travaux de l'équipe de W. Sebald (Duppatla *et al.* 2012), mais de petites modifications y ont toutefois été apportées.

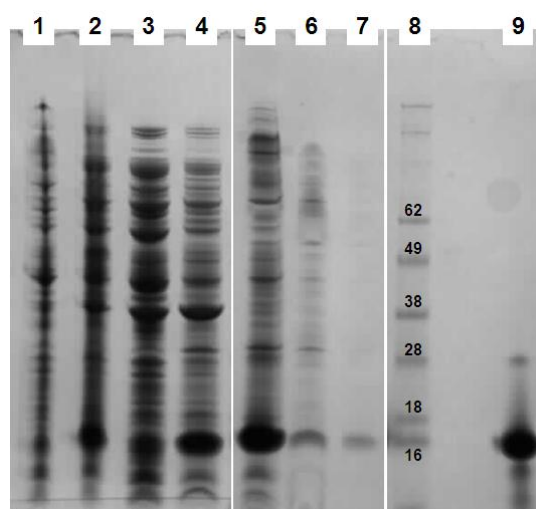
#### 3.1. Production de l'IL-4 mutée

##### 3.1.1. Production et purification

Pour établir une lignée cellulaire d'*E. coli* exprimant l'IL-4 mutée, l'équipe du Docteur Walter Sebald (Würzburg, Allemagne) a mis à notre disposition le plasmide correspondant.

La lignée recherchée a été obtenue grâce à une transformation par choc thermique. Pendant la période d'expression induite par l'IPTG (4 h environ), la cellule bactérienne

produit majoritairement l'IL-4 sous forme dénaturée dans des corps d'inclusion. Les parois et membranes cellulaires ont donc été détruites en utilisant un disrupteur de cellules puis les corps d'inclusion ont été récupérés par centrifugation et lavés plusieurs fois pour éliminer les protéines solubles. Les débris membranaires ont été solubilisés et éliminés par lavage avec une solution contenant du Triton X-100. La **Figure II.12.** montre le profil électrophorétique des fractions solubles obtenues après chaque étape de lavage. Notons que la protéine (15 kDa) est également présente sous forme soluble. Ceci explique pourquoi nous perdons, à chaque étape de lavage, une fraction de la protéine produite (ligne 3, 4 et 5 de la **Figure II.12.**).

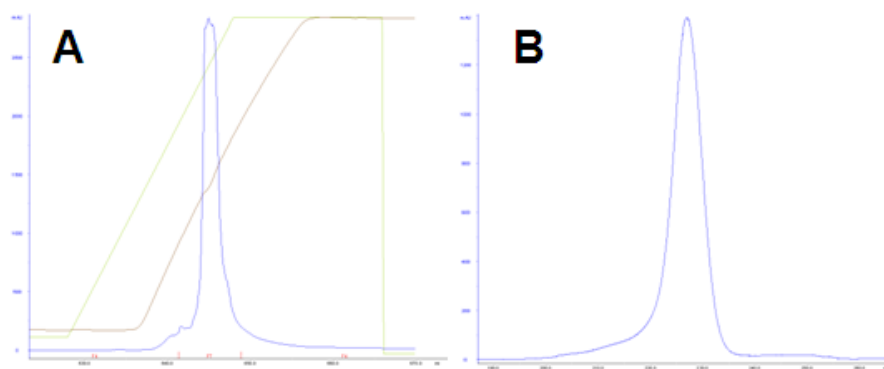


**Figure II.12. : Expression et purification de l'IL-4 mutée suivie par SDS-PAGE.**

**1 : Extrait cellulaire total avant induction. 2 : Extrait cellulaire total après induction. 3-7 : Surnageants issus des 5 étapes de lavage. 8 : Echelle de poids moléculaire en kDa. 9 : IL-4 mutée portant un glutathion après purification.**

La dénaturation des protéines contenues dans les corps d'inclusion s'effectue avec du chlorure de guanidinium. Pour promouvoir un repliement correct, les protéines sont placées dans une solution contenant de l'arginine, de l'EDTA et un mélange de glutathion oxydé/réduit (5/2). Après dialyse contre du PBS, la solution protéique est alors purifiée sur une colonne échangeuse de cation. La protéine éluée (**Figure II.13.A**), subit alors une étape de purification par perméation de gel (**Figure II.13.B**). L'IL-4 mutée purifiée porte alors sur le résidu muté S16C un glutathion provenant de l'étape de renaturation.

Les protocoles utilisés pour l'ensemble des étapes précédentes sont explicités en détail aux paragraphes 1.5. et 2.3, du chapitre IV. Toutes les étapes d'extraction et de purification sont conduites à 4°C. Pour obtenir environ 50 mg de protéine purifiée, il est nécessaire d'utiliser 10 L de culture cellulaire.



**Figure II.13. : Profil d'élution des colonnes de purification de l'IL-4.**

**A :** Elugramme de la colonne échangeuse de cation monoS. En bleu l'absorbance à 280 nm, en vert le pourcentage de la solution d'élution et en marron la conductance. **B :** Profil d'élution de la colonne d'exclusion de taille avec en bleu l'absorbance à 280 nm.

### 3.1.2. Déglutathionylation enzymatique

Le glutathion porté par l'IL-4 produite doit être retiré pour ancrer un foldamère. Pour cela, divers essais préalables de déglutathionylation chimique ont été réalisés en utilisant des réducteurs tels que le TCEP. Le potentiel redox des ponts disulfures naturellement présents dans la protéine étant très proche de celui existant entre la cystéine C16 et le groupement glutathion, la réduction contrôlée n'a malheureusement pas conduit à une réduction spécifique de ce pont disulfure.

Une méthode de réduction par voie enzymatique a donc été choisie. Pour cela, nous avons utilisé la glutarédoxine et la glutathion réductase. La réaction est suivie par la décroissance de l'absorption à 340 nm due à la disparition de NADPH (Duppatla *et al.* 2012). Une fois la réaction effectuée, le mélange est déposé sur une colonne d'exclusion de taille pour éliminer les enzymes présentes, les substrats restants et les produits non souhaités de la réaction. Avant de stocker la protéine concentrée, nous nous assurons qu'une seule cystéine est libre en utilisant le réactif et protocole d'Ellman (voir le paragraphe 2.5. du Chapitre IV : Matériels et méthodes et la référence Ellman, 1988). Les résultats obtenus ont été confirmés par spectrométrie de masse (ionisation en mode positif) et sont reportés dans le **Tableau II.2**. Les différences observées entre les masses théoriques et observées ne sont pas significatives au regard de la précision des appareils utilisés. Les résultats confirment que le glutathion a bien été retiré de l'IL-4 S16C.



	Masse théorique	Masse observée
IL-4S16C glutathion après purification	15 409	15 409
Radical glutathion	305	306
IL-4 S16C déglutathionylée	15 105	15 106

Tableau II.2. : Résultats des expériences de spectrométrie de masse en daltons.

## 3.2. Ligation de foldamères sur l'IL-4 S16C

A ce stade, la protéine mutée est prête à interagir avec les composés **A** ou **B**.

### 3.2.1. Ligation du composé **A** portant une fonction maléimide

Le nombre d'équivalents du composé **A** et la nature du tampon, utilisé à une concentration de 100 mM, ont été optimisés pour une concentration en protéine constante, égale à 133  $\mu$ M (soit 2 mg/mL). Les résultats, analysés par spectrométrie de masse (ionisation par électrospray en mode positif, analyseur à temps de vol), sont résumés dans le **Tableau II.3**. Les spectres correspondants, enregistrés à Slough (UCB Pharma) pendant ma seconde année de thèse, sont présentés dans la **Figure II.13**.

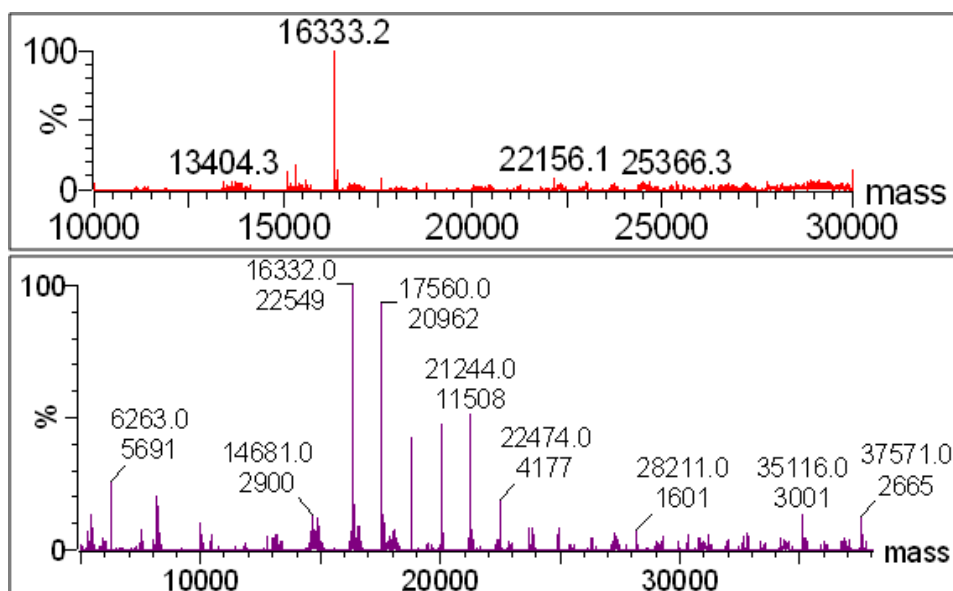
	1 équivalent	2 équivalents	3 équivalents	5 équivalents
Bis-tris pH 6,5	16 333	16 332 et 17 560 (21 244, 20 002, 18 780, 22 474)		
Hepes pH 7,0			16 332 et 17 560 (20 010, 21 244, 18 780, 22 474)	
Hepes pH 7,5		16 333 (traces d'autres pics)		16 332 et 17 559 (20 016, 21 248, 22 474, 18 780)

Tableau II.3. Résultats des expériences d'ancrage du composé **A** sur l'IL-4 mutée.

Les masses des pics secondaires sont rapportées entre parenthèses par ordre décroissant d'intensité. Les spectres des conditions en rouge et en bleu sont donnés dans la **Figure II.14**. Les masses sont données en daltons.

A la vue du tableau précédent, il s'avère que l'utilisation d'un tampon Bis-tris à pH 6,5 et la présence de 1 équivalent de composé **A** constitue la meilleure condition testée. La seule entité présente dans le milieu réactionnel possède en effet une masse correspondant à la masse théorique de l'IL-4 mutée portant une molécule **A** ( $15\,105 + 1\,228 = 16\,333$  Da) (**Figure II.14.** spectre rouge). Un signal très minoritaire apparaît néanmoins et pourrait correspondre à la protéine n'ayant pas réagi. Ce signal étant réellement très faible par rapport au précédent, nous pensons que la réaction est quasiment totale, bien que nous sachions que la spectrométrie de masse, utilisée dans de telles conditions expérimentales, ne puisse être considérée comme une méthode d'analyse quantitative.

Si dans le même tampon utilisé au même pH, 2 équivalents du composé **A** sont présents (**Figure II.14.** spectre violet), le pic de plus haute intensité correspond toujours à l'adduit IL-4 – A (16 332 Da). Néanmoins, d'autres pics apparaissent aux côtés de ce dernier. Ils correspondent à la masse de l'IL-4 mutée portant 2, 3, 4, 5 voire 6 molécules A (masses observées : 17 560, 18 780, 20 002, 21 244, 22 474 Da, pour des masses théoriques respectives de 17 561, 18 786, 20 014, 21 242, 22 474 Da). Ce résultat n'est pas surprenant. En règle générale, la fonction maléimide ne réagit qu'avec les fonctions thiols des protéines mais il arrive que lorsque le pH ou que les concentrations en réactif sont suffisamment élevés, des réactions avec les fonctions amines et/ou hydroxyles des protéines soient possibles.



**Figure II.14.** Spectres obtenus après réaction d'ancrage du composé **A**.

Tampon Bis-tris à pH 6,5 avec 1 (spectre rouge) ou 2 (spectre violet) équivalents du composé **A**. Les masses rapportées sont en daltons.

### 3.2.2. Ligation du composé **B**, portant une fonction thiol activée

Nous avons réalisé le même type d'expériences (mêmes concentrations en protéine et en tampon) avec le composé **B**. Chaque mélange réactionnel a été analysé par spectrométrie de masse de la même manière que précédemment. Les résultats sont reportés dans le **Tableau II.4**.

	1 équivalent	2 équivalents	5 équivalents	10 équivalents
Cit Na pH 4,0				15 104 (16 210)
Cit Na pH 5,0		15 105 (18 425, 17 319,16 212,)		
Hepes pH 7,0	15 105 et 16 210			16 211 (15 210)
Hepes pH 8,0			18 425 (16 211, 17 429)	
Glycine pH 9,0	15 103 et 16 209 (18 424, 15 213)	16 210 (18 425, 15 214)	16 210 (18 425, 15 214)	18 426, 16 211 et 17 429 (15 214, 16 431 et 18 524)

**Tableau II.4.** Résultats des expériences d'ancrage du composé **B** sur l'IL-4 mutée.

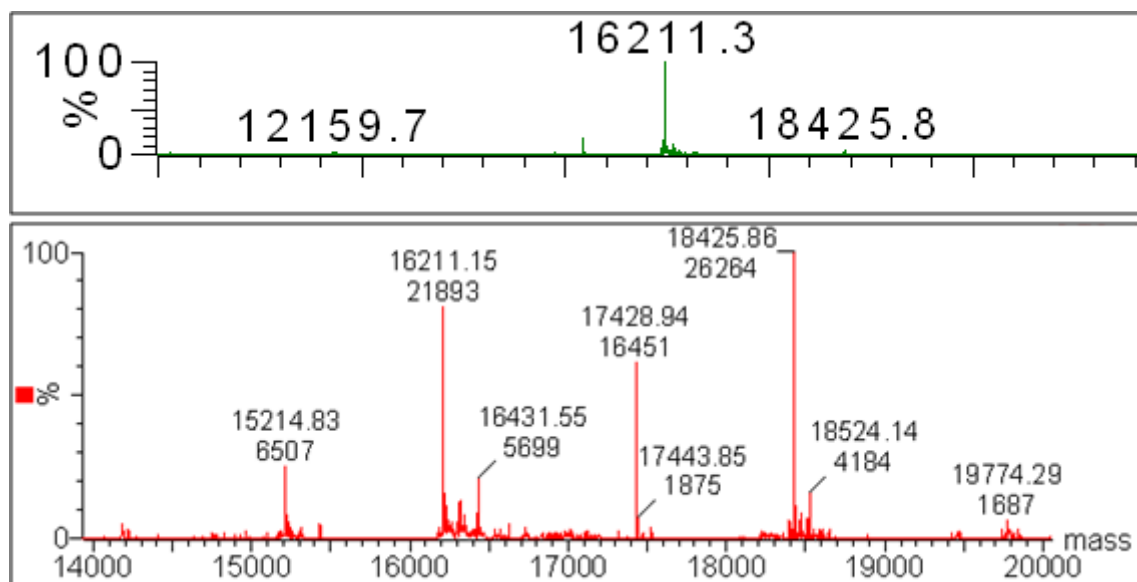
Les masses des pics secondaires sont rapportées entre parenthèses par ordre décroissant d'intensité. Les spectres des conditions en rouge et en vert sont donnés dans la Figure II.15. Les masses sont exprimées en daltons.

L'augmentation de masse correspondant à l'ajout d'une molécule **B** sur l'IL-4 mutée est de 1 105 Da. Cette masse correspond en fait à la masse du composé **B** (1 215 Da) auquel doit être retranchée la masse (110 Da) du groupe partant (voir schéma réactionnel de la Figure II.9).

La condition correspondant à 10 équivalents de composé **B** dans un tampon Hepes à pH 7,0, constitue la meilleure condition testée. En effet l'entité moléculaire associée au pic de plus haute intensité possède une masse (16211 Da) égale, aux incertitudes de mesure près, à celle attendue (masse théorique de l'IL-4 mutée portant un foldamère **B** : 15 105 + 1 105 = 16 210 Da) (**Figure II.15**, spectre vert). La réaction semble totale. Seul un signal de très faible intensité apparaît également. Il correspondrait à la protéine ayant réagi avec un groupe partant (masse théorique : 15 105 + 110 = 15 215 Da).

Une analyse similaire peut être conduite pour les réactions réalisées dans un tampon glycine à pH 9,0 en présence de 2 ou 5 équivalents du composé **B**. Les spectres obtenus à partir de ces conditions montrent un pic majoritaire à 16 210 Da (protéine ayant fixé une molécule **B**) et deux pics de faible intensité. Ces derniers correspondent respectivement aux masses attendues après addition de trois molécules **B** sur la protéine mutée ( $15\,105 + 3 \times 1\,105 = 18\,420$  Da) ou à l'ajout d'un groupe partant sur la protéine (15 215 Da).

La réaction qui s'effectue dans le tampon glycine à pH 9,0 avec 10 équivalents semble plus complexe. Dans le spectre correspondant (**Figure II.15**, spectre rouge) apparaissent de multiples pics. Un pic intense (18 426 Da) correspond à l'addition de 3 molécules **B** à la protéine mutée (masse théorique : 18 420 Da). Deux autres pics (16 211 et 17 428 Da) correspondent respectivement à l'adduit IL-4 - **B** et à la protéine mutée ayant fixée deux molécules **B** et un groupe partant ( $15\,105 + 2 \times 1\,105 + 110 = 17\,425$  Da). D'autres pics, de plus faible intensité, sont également présents. Ils peuvent être associés à l'IL-4 mutée ayant fixé respectivement un groupement partant (masse théorique : 15 215 Da), deux groupements partants et une molécule **B** (masse théorique :  $15\,105 + 2 \times 110 + 1\,105 = 16\,430$  Da), ou à la formation de divers autres complexes.



**Figure II.15.** Spectres obtenus après réaction d'ancrage du composé **B**.

Spectres obtenus en présence de 10 équivalents de composé **B** et dans le tampon Hepes à pH 7,0 (spectre vert) ou dans le tampon glycine pH 9,0 (spectre rouge). Les masses rapportées sont en daltons.

Les expériences de spectrométrie de masse précédentes soulèvent quelques questions dès lors que l'on observe, pour certaines conditions (le plus souvent correspondant à un excès

de foldamère), une addition non pas d'un mais de plusieurs composés **A** ou **B**. Nous avons émis une explication à ces observations si l'on considère le seul composé **A**. En cas d'excès de maléimide et à certains pH, de tels composés peuvent réagir avec des fonctions autres que des thiols. Le problème est plus délicat dès lors qu'il s'agit du composé **B** si l'on considère que ce dernier ne peut réagir qu'avec une fonction thiol. Si l'on exclut toute interaction non spécifique avec la protéine (interaction foldamère/protéine par exemple) difficile à observer en spectrométrie de masse, il faut considérer un nombre de cystéines libres supérieur à un et donc la réduction totale ou partielle d'un ou plusieurs ponts disulfures.

Les expériences précédentes montrent la faisabilité d'ancrer un unique foldamère à la protéine, quel que soit la fonction (maléimide ou thiol activé) utilisée. Afin d'éviter l'introduction d'un centre chiral, nous avons choisi de poursuivre ce travail uniquement avec le composé **B** greffé à la protéine. Restait toutefois à confirmer que dans le cas d'un seul composé fixé à la protéine, celui-ci se fixait bien sur la cystéine C16. Pour cela, des expériences de spectrométrie de masse sur des échantillons ayant subi une digestion trypsique (expériences réalisées en collaboration avec Jean-Marie Schmitter, CBMN, Bordeaux) ont été conduites. Les résultats sont présentés dans la **Figure II.16**.

Rappelons que 3 ponts disulfures (C3-C127, C46-C99 et C24-C65) sont présents dans la protéine native.

Avant la réaction d'ancrage du composé **B** à la surface de l'IL-4 mutée, la digestion à la trypsine fait apparaître un massif de pics. A l'un d'entre eux est associée une masse de 2 124 Da (**Figure II.16.A**) correspondant aux peptides 43-47 et 89-102 liés de façon covalente par le pont disulfure natif C46-C99.

Une fois la réaction d'ancrage réalisée, de nouveaux pics apparaissent, notamment à 2 155 Da et 2 337 Da. Ces pics correspondent respectivement aux peptides 13-21 (1 049 Da) et 65-75 (1 231 Da) portant chacun une molécule **B** (**Figure II.16.B**). Ainsi il semblerait que le foldamère ne soit pas ancré exclusivement sur la cystéine C16, mais aussi sur la cystéine C65. Ceci impliquerait que la protéine mutée ne soit pas toujours correctement repliée et que dans certains cas, ce soit la cystéine C65 qui reste libre.

Si le mélange réactionnel est préalablement dialysé avant d'être digéré puis analysé par spectrométrie de masse, les pics correspondants à l'addition du composé **B** sur la cystéine C65 ont une intensité réduite par rapport à ceux observés en absence de dialyse (**Figure II.16.C**). L'intensité des autres pics est par contre conservée. La membrane de

dialyse utilisée ayant un *cut-off* de 10 kDa proche du poids moléculaire de la protéine (15 kDa), il semble possible que lorsque la protéine n'est pas correctement repliée, cette dernière puisse passer à travers les pores. Cette simple observation montre qu'il serait alors possible de séparer les deux populations.

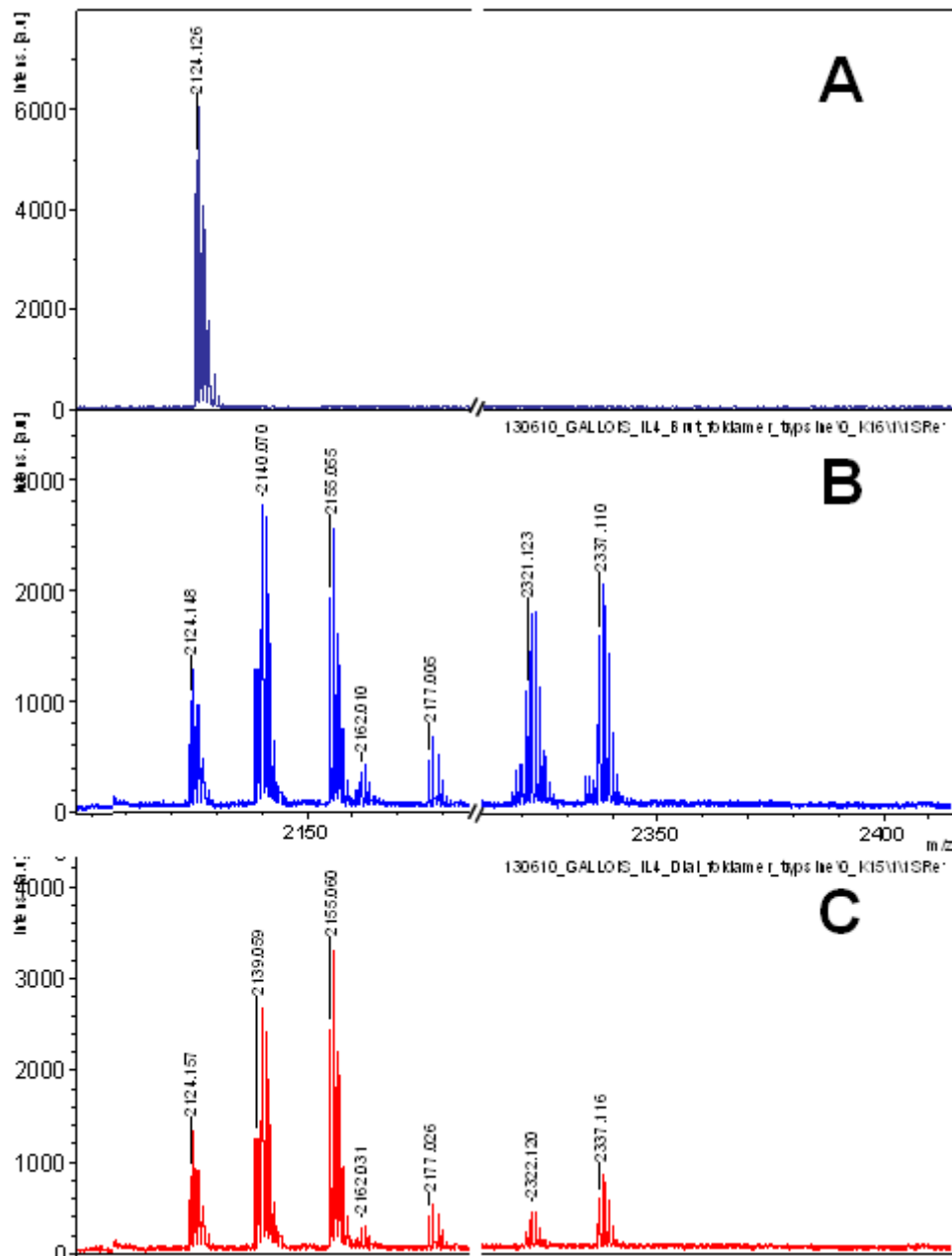
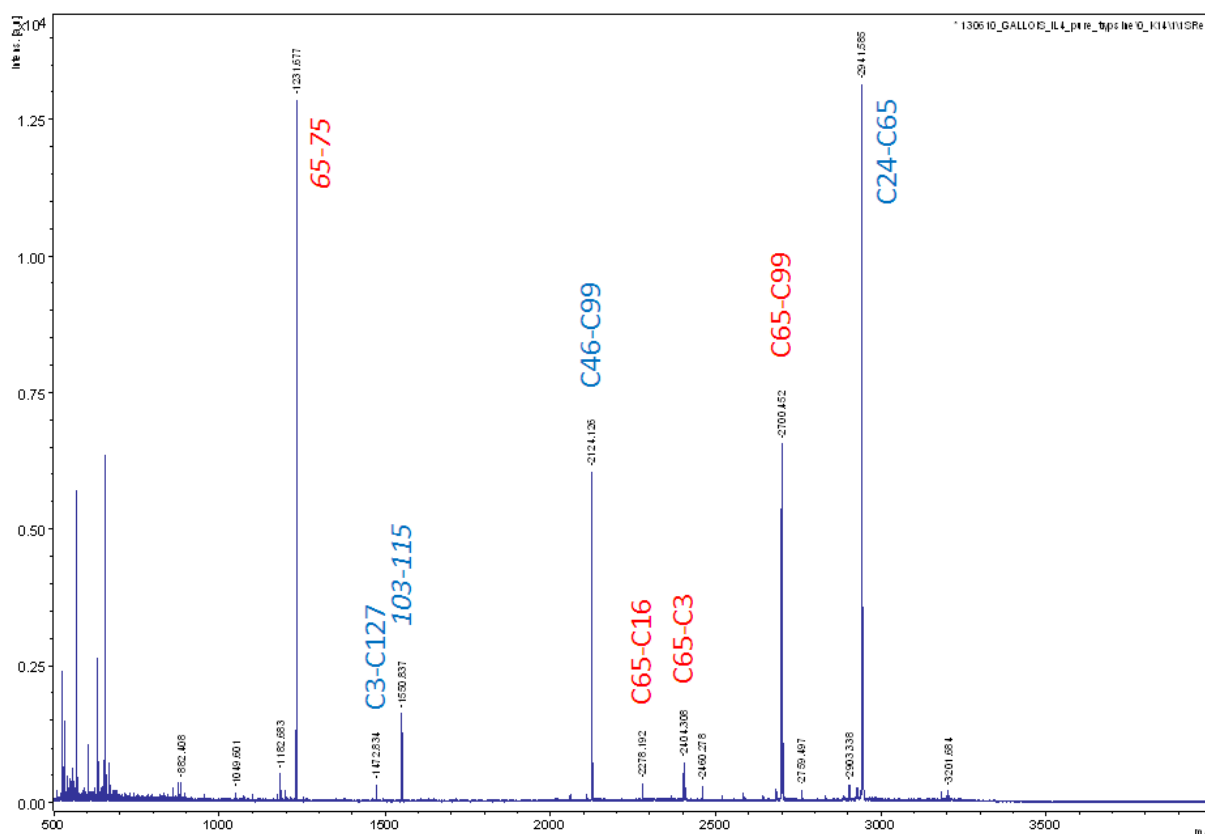


Figure II.16. : Spectres de masse après digestion à la trypsine.

A : Protéine mutée. B : Protéine mutée après ancrage. C : Protéine mutée après ancrage et dialyse. Les masses indiquées sont en daltons.

### 3.2.3. Mauvais repliement de l'IL-4

Le spectre de masse de la digestion de l'IL-4 mutée seule (**Figure II.17.**) nous indique la présence du peptide 103-115 (1 550 Da) et de fragments (masses expérimentales : 1 472, 2 124 et 2 941 Da pour des masses théoriques respectives de 1 470, 2 126, 2 944 Da) incluant les 3 ponts disulfures natifs. Trois autres fragments sont présents (masses expérimentales : 2 404, 2 278 et 2 700 Da pour des masses théoriques respectives de 2 406, 2 280 et 2 703 Da). Chacun d'eux inclut 1 nouveau pont disulfure (C65-C3, C65-C16 et C65-C99). En outre, ces données montrent la présence du peptide 65-75 (masse expérimentale : 1 231 Da pour une masse théorique de 1 231 Da) montrant de façon indiscutable la présence de la cystéine C65 sous forme réduite.



**Figure II.17. : Détail du spectre de masse de la digestion de l'IL-4 mutée.**

Les pics relatifs à la protéine correctement repliée sont en bleu, ceux correspondant à un mauvais repliement sont en rouge. Les masses sont exprimées en daltons.

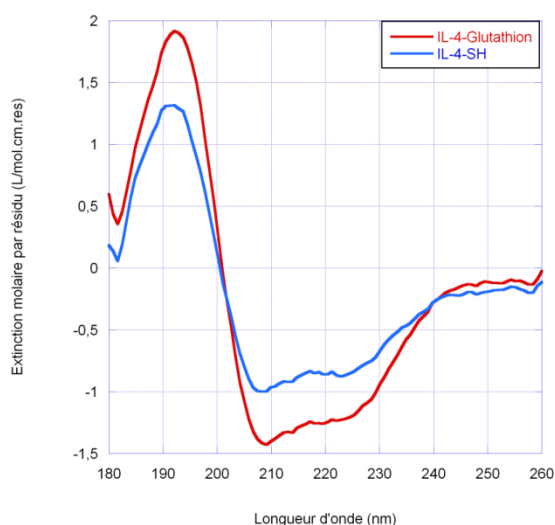
L'ensemble de ces résultats tendrait à démontrer que l'IL-4 mutée qui a été utilisée n'est pas, dans sa totalité, correctement repliée. Ce mauvais repliement peut se produire à

différentes étapes, notamment pendant les phases de renaturation et de purification, ou pendant la réaction de déglutathionylation.

### 3.2.4. Conclusion et discussion des résultats obtenus

Des tentatives de cristallisation de l'IL-4 mutée portant un glutathion ont été réalisées. Elles ont permis d'obtenir des cristaux par la technique de diffusion de vapeur en gouttes suspendues en utilisant une solution de protéine à 662  $\mu$ M ( $\sim 10$  mg/mL) dans 50 mM de Tris pH 8,0. Le réservoir contenait 2,4 M  $(\text{NH}_4)_2\text{SO}_4$  dans 40 mM de cacodylate de sodium, pH 6,0. Les cristaux sont apparus après trois semaines environ et ont été cryo-protégés par trempage dans une solution de réservoir complétée par 23 % de glycérol. Ces cristaux ont été testés sur la ligne de lumière Proxima 1 à Soleil (Gif sur Yvette). Un jeu de données a pu être collecté et la structure résolue par remplacement moléculaire en utilisant comme modèle la structure de l'IL-4 native à 1,8 Å de résolution (code PDB : 2B8U). La structure trouvée (données non publiées) montre la présence d'un glutathion partiellement résolu, lié à la cystéine C16. Cette structure ne diffère que très peu du modèle de référence puisque le r.m.s.d. calculé à partir de la position de 127 C $\alpha$  superposables est de 0,24 Å (écart de position maximal de 1,25 Å pour les résidus Q20). Ces résultats confortent la présence en solution d'une quantité non négligeable de protéine correctement repliée.

De plus nous avons mesuré les spectres CD de la protéine après purification (portant alors un glutathion) et après la réaction de déglutathionylation (**Figure II.18.**).



**Figure II.18. : Spectres CD de l'IL-4 mutée avant et après déglutathionylation.**

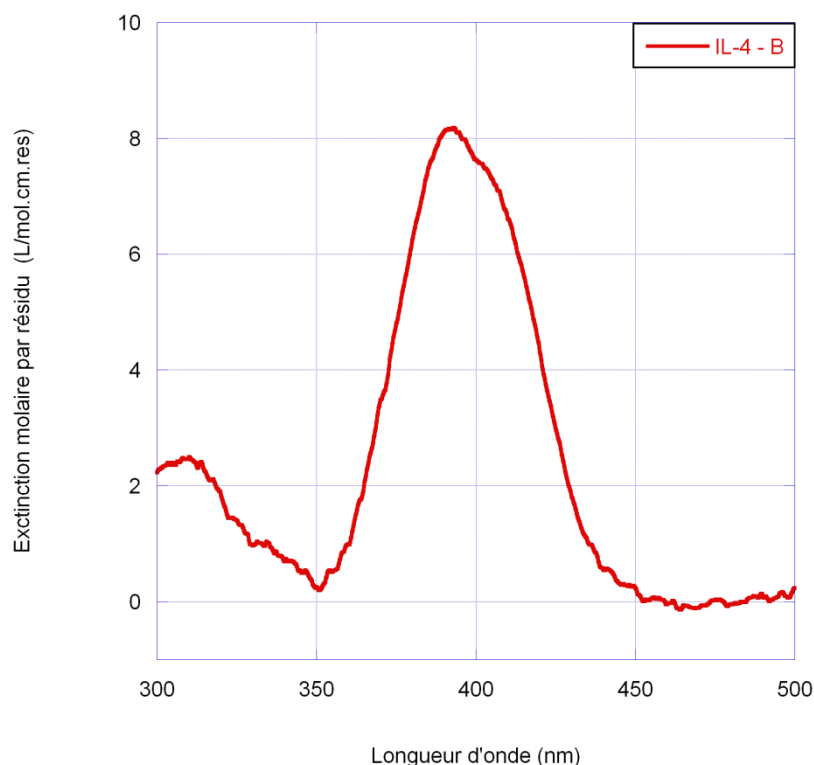


La comparaison de ces deux spectres ne montre pas de changement conformationnel notable, ce qui tend à prouver que le pourcentage relatif des éléments de structure secondaire est conservé avant et après déglutathionylation.

A la vue des résultats précédents, tout porte à croire que l'étape de dénaturation/renaturation permet l'obtention d'une quantité significative de protéine mutée correctement repliée, et que la réaction de déglutathionylation n'affecte pas notablement le repliement. Le stockage de la protéine à -20°C pendant une période de quelques mois pourrait être à l'origine des différences de repliement observées après digestion trypsique. Répondre à cette question impliquerait de faire la digestion trypsique sur un lot de protéine fraîchement purifiée. Les impératifs du projet de recherche (mobilité, disponibilité des appareillages) n'ont pas permis de mener à bien cette étude au moment voulu. Ceci nous apparaît dommageable. Si le stockage de la protéine n'affecte pas sa stabilité, cela sous-entend que la méthode de purification décrite dans ce rapport (extraction de la protéine à partir des corps d'inclusion), et préalablement utilisée par le groupe de Walter Sebald à Würzburg (Allemagne) dans diverses études, n'est peut-être pas la plus appropriée. Il serait sans doute alors préférable de purifier la protéine soluble lors de l'expression. Dans ce cas, soulignons toutefois que les rendements seraient encore moins performants que ceux, déjà limités, obtenus à partir des corps d'inclusion.

### 3.3. Derniers résultats obtenus

En dépit des difficultés rencontrées, nous avons souhaité savoir si le foldamère **B**, ancré à la protéine, interagissait avec la protéine. Pour cela nous avons donc mesuré à 25 °C un spectre de dichroïsme circulaire induit sur l'adduit obtenu. Une solution de ce dernier à une concentration de 66  $\mu$ M dans un tampon NaH<sub>2</sub>PO<sub>4</sub> 50 mM pH 7,5 a été utilisée. Le spectre obtenu après ancrage et purification est présenté **Figure II.18**. L'interaction entre un foldamère et une protéine peut être stéréosélective et donc modifier l'équilibre existant entre les deux conformères d'hélicité opposée. Si c'est le cas, un signal de dichroïsme circulaire induit pourra être observé (voir Partie IV. Matériels et méthodes paragraphe 3.1.).



**Figure II.18. : Spectre ICD de l'adduit IL-4-B obtenu à 25 °C.**

Ce spectre met en évidence un pic d'absorption à 390 nm. Ce signal positif, d'intensité notable, est attribué d'après les résultats publiés par Kendhale *et al.* (2011), à la prépondérance d'une hélicité droite du foldamère. Le spectre du composé **B** seul ne révèle quant à lui, aucun pic dans cette gamme de longueurs d'onde (données non reportées ici).

Ce nouveau résultat est d'importance. Il permet d'affirmer que la protéine induit une chiralité préférentielle au foldamère qui lui est attaché. Des interactions existeraient donc entre partenaires, bien qu'aucune optimisation de la séquence du foldamère n'ait été réalisée.

## Conclusion

Malgré les difficultés rencontrées, nous sommes parvenus à montrer la faisabilité de notre hypothèse de travail : un foldamère de quinolines est capable d'acquérir une hélicité préférentielle lorsqu'il "voit" une protéine. Il interagit donc vraisemblablement avec cette dernière de façon spécifique.

Pour parvenir à ce résultat, nous avons identifié des conditions d'ancrage d'un foldamère au voisinage de la surface protéique à reconnaître, et valider le choix de l'interleukine 4 comme protéine modèle.

Ce résultat est pour le moins encourageant sachant que la fonctionnalisation du foldamère n'a pas été optimisée à ce stade de notre travail. Il ne nous a pas été possible, dans le temps imparti, de finaliser cette étude, c'est-à-dire de caractériser les interactions induisant le changement d'hélicité. Cela nécessitera la cristallisation de cet adduit et la résolution de la structure par diffraction des rayons X. Nous aurions aimé présenter cette étude. Nous sommes toutefois conscients que pour y parvenir des quantités d'adduit plus importantes devront être obtenues. Cela implique de résoudre le problème concernant le repliement partiellement incorrect de la protéine.

### Chapitre III

L'anhydrase carbonique humaine, une preuve  
de concept



## Introduction

L'inhibition des interactions protéine-protéine par des foldamères aromatiques fonctionnalisés est un projet particulièrement ambitieux reposant pour l'instant sur très peu de données expérimentales. La reconnaissance d'un tel concept nécessite d'en démontrer la faisabilité. Une étape importante sera franchie dès lors qu'il aura pu être montré qu'un foldamère peut reconnaître une surface protéique cible, préalablement définie. Afin de simplifier la démarche expérimentale, nous avons recherché une protéine facile à produire, à purifier et à cristalliser, pour laquelle les propriétés tant structurales que fonctionnelles étaient parfaitement déterminées. L'anhydrase carbonique humaine (HCA-II) est très vite apparue être une protéine de choix. Cette enzyme est en effet un système modèle très largement utilisé comme référence dans l'étude des propriétés de liaison d'un ligand à une enzyme (Krishnamurthy *et al.* 2008).

# 1. L'anhydrase carbonique, une protéine modèle

---

## 1.1. Description et choix de la protéine

Les anhydrases carboniques (EC 4.2.1.1) sont des métalloenzymes à zinc de poids moléculaire voisin de 30 kDa. Ces enzymes sont ubiquitaires. Elles ont été étudiées dans divers microorganismes, plantes, et animaux. Il existe 5 familles d'anhydrases carboniques. Deux d'entre elles,  $\delta$  et  $\epsilon$ , n'ont été que récemment découvertes (Maupin *et al.* 2009). Les trois autres ( $\alpha$ ,  $\beta$ , et  $\gamma$ ) sont connues depuis plus longtemps et leurs propriétés structurales sont parfaitement caractérisées. Actuellement 573 structures cristallographiques relatives aux anhydrases carboniques sont disponibles à la Protein Data Bank (PDB).

La fonction de ces enzymes est de catalyser, *in vivo*, l'hydratation réversible du dioxyde de carbone ( $\text{CO}_2$ ) en ion hydrogénocarbonate (ou bicarbonate  $\text{HCO}_3^-$ ) en relâchant un proton, par un mécanisme ping-pong (**Figure III.1.**). Outre cette activité d'hydratation du  $\text{CO}_2$ , elles sont capables de catalyser des réactions d'hydratation des aldéhydes aliphatiques et

aromatiques (Shrerdan *et al.* 1980), et d'hydrolyser *in vitro* une large variété d'esters (Pocker et Sarkanen 1978).

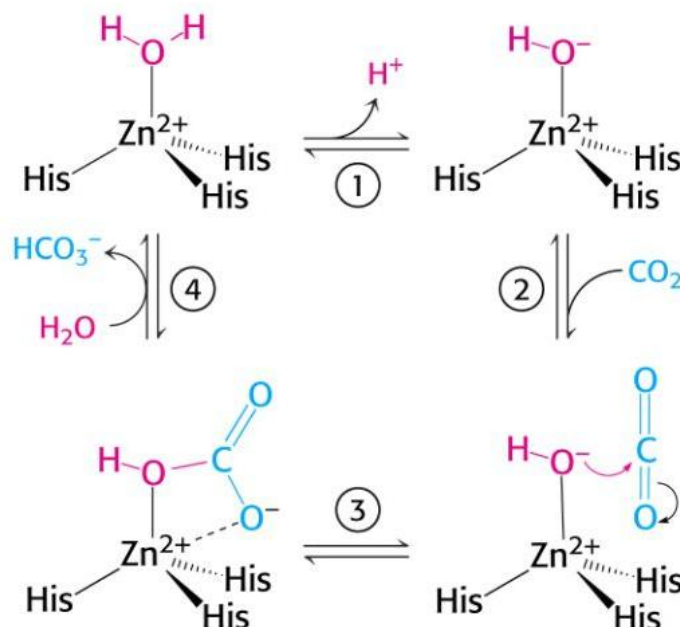


Figure III.1. : Schéma réactionnel enzymatique de l'hydratation du dioxyde de carbone.

1 : Le zinc abaisse le pKa de la molécule d'eau qui lui est ligandée de sorte qu'elle perde un proton. 2 : Le  $\text{CO}_2$  se fixe au site actif. 3 : L'ion hydroxyde est suffisamment nucléophile pour attaquer le  $\text{CO}_2$  qui devient l'ion bicarbonate. 4 : Le site catalytique est régénéré en libérant  $\text{HCO}_3^-$  et en fixant une nouvelle molécule d'eau.

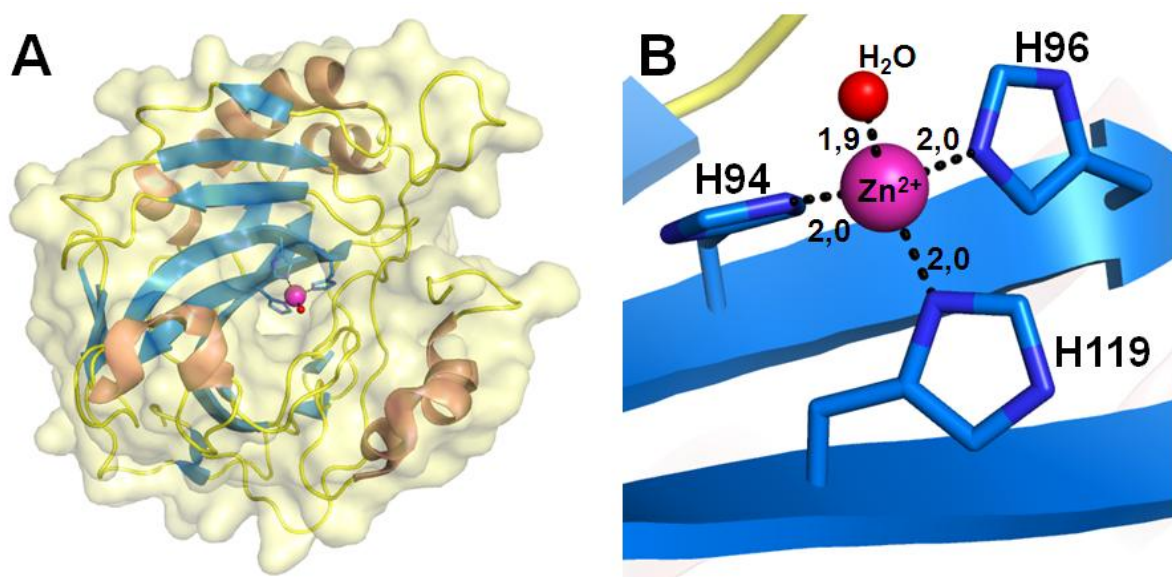
Leur capacité à hydrolyser le  $\text{CO}_2$ , à une vitesse limitée par la seule diffusion des molécules, rend ces enzymes centrales dans la respiration (transport du  $\text{CO}_2$ ), la photosynthèse, les processus sécrétoires et le maintien du pH cellulaire entre autres (Maupin *et al.* 2009).

Avec une constante catalytique d'environ  $10^6 \text{ s}^{-1}$  (Maupin 2009), la HCA-II est l'isoenzyme la plus rapide des anhydrases carboniques  $\alpha$ . Elle constitue une cible thérapeutique de la chimie médicinale actuelle, et son inhibition par de petites molécules porteuses d'une fonction sulfonamide aromatique est importante dans le traitement du glaucome, de l'épilepsie et de certains cancers (Weber *et al.* 2004 - Nishimori *et al.* 2007).

La structure tridimensionnelle de la HCA-II a été déterminée par diffraction aux rayons X pour la première fois en 1972 (Liljas *et al.* 1972), à la résolution de 2,0 Å. Elle a été

maintes fois redéterminée depuis sous diverses formes (native, mutée, en complexe ou non avec de très nombreux ligands, ...). Avec 468 structures déposées, la HCA II est de loin la plus étudiée de toutes les anhydrases carboniques. Depuis quelques années, nous disposons d'une structure déterminée à résolution atomique (résolution : 0,9 Å ; code PDB : 3KS3) (Avvaru *et al.* 2010).

C'est une protéine globulaire, constituée d'une seule chaîne protéique (260 acides aminés). Sa structure tertiaire est constituée d'un feuillet central de 10 brins  $\beta$  entouré de 7 hélices  $\alpha$  (**Figure III.2.A**). Le site actif est formé d'une cavité amphiphile dont l'ouverture vers l'extérieur et la profondeur sont sensiblement de 15 Å. Au fond de cette cavité, se trouve un cation  $\text{Zn}^{2+}$ , jouant le rôle de cofacteur. L'ion  $\text{Zn}^{2+}$  se coordonne selon une géométrie tétraédrique aux trois fonctions imidazoles de trois résidus histidine, H94, H96 et H119, ainsi qu'à un quatrième ligand, une molécule d'eau (**Figure III.2.B**).

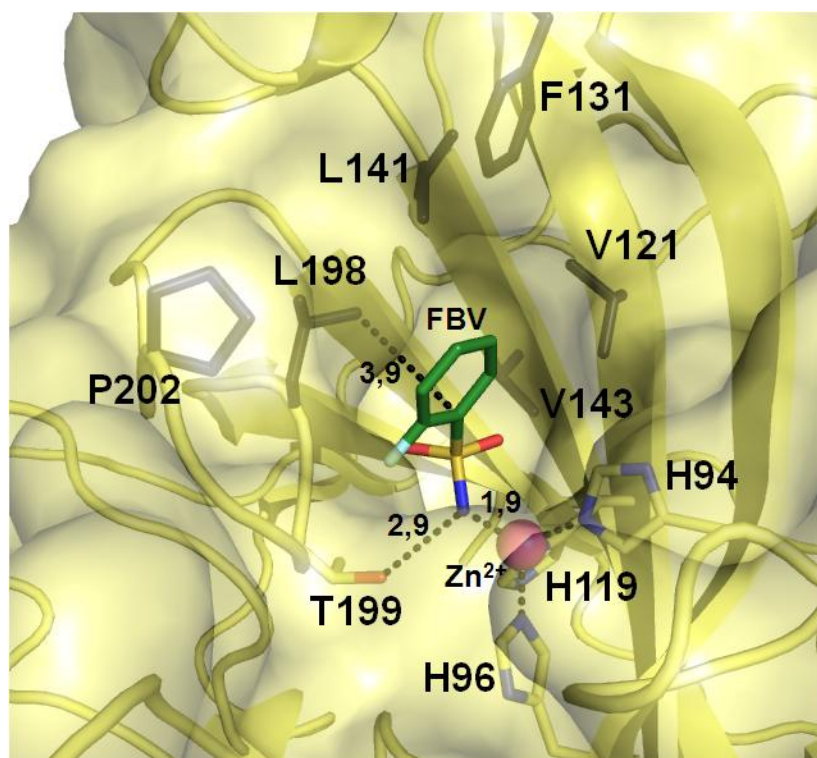


**Figure III.2. : Structure générale et site actif de la HCA-II.**

**A :** La protéine est matérialisée par sa surface et par une représentation en ruban permettant de voir les structures secondaires (hélices  $\alpha$  en rouge, brins  $\beta$  en bleu et boucles en jaune). **B :** Site actif où le zinc est coordonné à trois histidines et une molécule d'eau. Les distances sont données en Å.

Les inhibiteurs de cette enzyme, des sulfonamides aromatiques substitués, interagissent avec la protéine par l'intermédiaire de différentes interactions : complexation de la fonction sulfonamide avec le zinc ; liaisons hydrogène en particulier avec T199 ; interactions hydrophobes avec la face apolaire du site actif et en particulier avec L198 (**Figure III.3.**)





**Figure III.3. : Interactions entre l'inhibiteur FBV (2-fluorobenzensulfonamide) et la HCA-II.**

L'inhibiteur est « posé » sur la moitié hydrophobe de la cavité. FBV interagit grâce à l'atome d'azote de la fonction sulfonamide par une liaison hydrogène avec T199 et crée une liaison de coordination avec le zinc du site actif. Les distances sont données en Å (code PDB : 2WEG).

Les méthodes permettant d'isoler et de purifier l'enzyme sont très largement décrites dans la littérature. Ainsi, les industriels isolent préférentiellement la protéine à partir d'érythrocytes dans lesquels l'anhydrase carbonique est très abondante. D'autres méthodes, basées sur la biologie moléculaire et différents systèmes d'expression, autorisent la production de l'enzyme sous forme recombinante (Banerjee *et al.* 2004). La purification s'effectue le plus souvent sur colonne d'affinité, modifiée par greffage d'un groupement sulfonamide. Cette technique, relativement simple, assure une pureté supérieure à 90 %, autorisant de très nombreuses études (dont des études structurales) sans envisager d'étapes de purification supplémentaires. Après purification, la protéine reste stable au cours du temps.

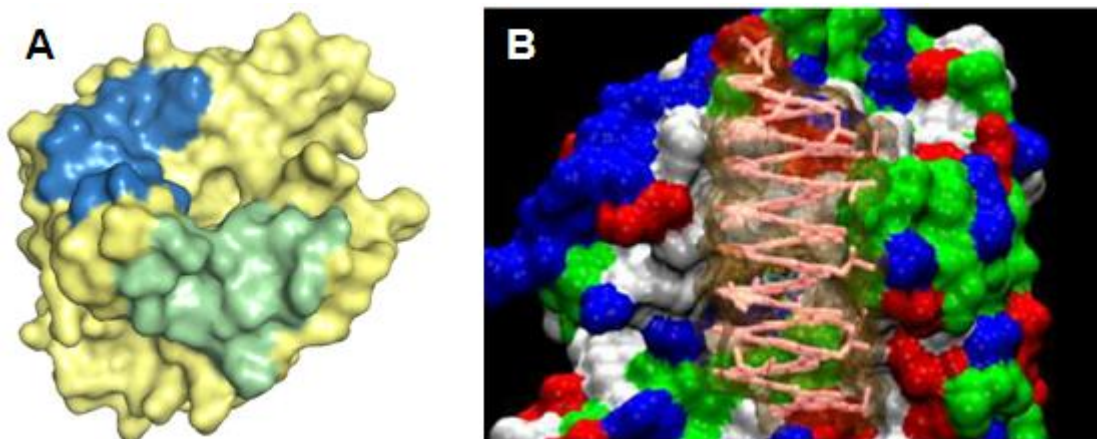
Les conditions de cristallisation de l'enzyme sont parfaitement connues et font appel à des méthodes classiques fréquemment utilisées dans le cas des protéines (diffusion de vapeur en gouttes suspendues ou en gouttes assises). La HCA-II cristallise en présence de sulfate d'ammonium, de chlorure de sodium et de *para*-hydroxymercuribenzoate (pHMB) pour des valeurs de pH le plus souvent comprises entre 4,7 et 10,0 (Krishnamurthy *et al.* 2008). Le

groupe d'espace le plus fréquemment rencontré est le groupe d'espace monoclinique  $P2_1$  (paramètres de maille de la structure la plus résolutive (3KS3) :  $a = 42,35 \text{ \AA}$ ,  $b = 41,36 \text{ \AA}$ ,  $c = 72,31 \text{ \AA}$ ,  $\alpha = \gamma = 90,00^\circ$  et  $\beta = 104,42^\circ$ ).

## 1.2. Stratégie appliquée pour la reconnaissance de surface de la HCA-II par un foldamère

### 1.2.1. Stratégie générale

La structure tridimensionnelle de HCA-II étant connue, il est possible, à l'aide de différents logiciels graphiques, de déterminer son enveloppe de surface. Cette dernière montre, au voisinage du site actif, la présence de sillons, formes complémentaires de celle d'une hélice, dans lesquels pourrait s'insérer un foldamère (**Figure III.4.**). Des études préliminaires de modélisation moléculaire, menées dans le groupe de Michel Laguerre (IECB, Bordeaux) à partir des structures cristallographiques de la HCA-II et de divers foldamères, confirment cette possibilité. Les principaux résidus présents à la surface de ces sillons ont été détaillés dans le **Tableau III.1.**



**Figure III.4. : Deux sillons présents à la surface de la protéine pouvant accueillir un foldamère.**

**A : Les deux sillons principaux de la protéine au voisinage du site actif sont visualisés en bleu (sillon 1) et vert (sillon 2). B : Modélisation d'un foldamère dans le sillon 1. La surface de la protéine est colorée en bleu, rouge, vert et blanc en fonction de la nature de ses résidus (respectivement chargés positivement, négativement, polaires non chargés, et hydrophobes)**

Principaux résidus du sillon 1 (bleu)	Principaux résidus du sillon 2 (vert)
L57, R58, E69, F70, D71, D72, S73, R89, I91, Q92, W123, F127, K129, D130, F131, K133	D19, F20, P21, I22, G25, E26, R27, Q28, V135, Q136, Q137, P138, P201, P202, L204, E205, C206

**Tableau III.1. : Principaux résidus présents à la surface des sillons de la HCA-II.**

Pour obtenir une reconnaissance spécifique et de très haute affinité, il serait nécessaire de fonctionnaliser le foldamère de façon telle que certaines chaînes latérales des unités quinoline présentent une complémentarité de fonction avec les chaînes latérales des acides aminés se trouvant à la surface de la protéine. Pour y parvenir, il conviendrait d'utiliser une technique de synthèse linéaire, longue et fastidieuse, mais permettant à chaque étape, d'allonger le foldamère d'une unité monomérique différente des précédentes.

Cribler des séquences entières de foldamères dans l'espoir que l'une d'elles présente une bonne affinité pour la cible nécessiterait de très nombreuses synthèses. Les chimistes du groupe d'Ivan Huc ont donc opté pour une approche différente. Cette dernière utilise un système d'ancrage qui maintient le foldamère au voisinage de la surface protéique, tout en lui assurant un certain degré de mobilité. La méthodologie dite du « *structure based design* » s'appuie sur des données structurales, et permet d'optimiser et d'augmenter le nombre d'interactions entre les chaînes latérales du foldamère et la surface de la protéine cible. Cette stratégie implique donc de développer l'hélice étape par étape.

Le système d'ancrage retenu est basé sur la très forte affinité d'un des inhibiteurs naturels pour le site actif de la HCA-II. Il utilise un espaceur, liant un inhibiteur modifié au foldamère d'intérêt (**Figure III.5.**). L'optimisation de la longueur et de la nature de l'espaceur sera également réalisée *in silico* par modélisation moléculaire. L'objectif, à terme, est de développer un foldamère présentant une affinité suffisamment importante pour la surface cible, permettant de s'affranchir du système d'ancrage et d'observer une reconnaissance de surface spécifique par simple mise en solution du foldamère avec la protéine.

L'inhibiteur choisi est un arylsulfonamide, le N-benzyle-4-sulfamoylbenzamide (SBB), présentant une très forte affinité pour la HCA-II ( $K_D = 1,1$  nM) (Jain *et al.* 1994 - Abbate *et al.* 2002). La structure cristallographique du complexe SBB/HCA-II (code PDB = 1G4O) montre que l'extrémité benzyle terminale de l'inhibiteur affleure à la surface de la protéine et à proximité de ses sillons (**Figure III.5.** et **Figure III.6.**).

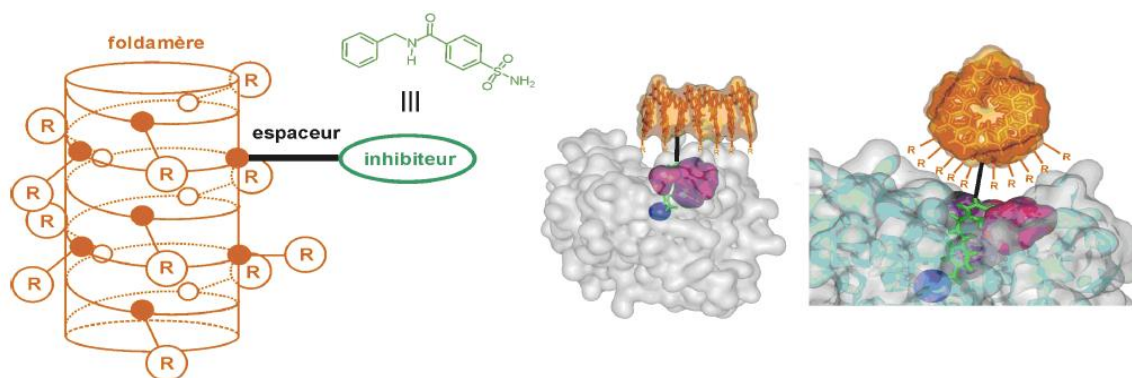


Figure III.5. : Stratégie de reconnaissance de surface de la HCA-II par un foldamère.

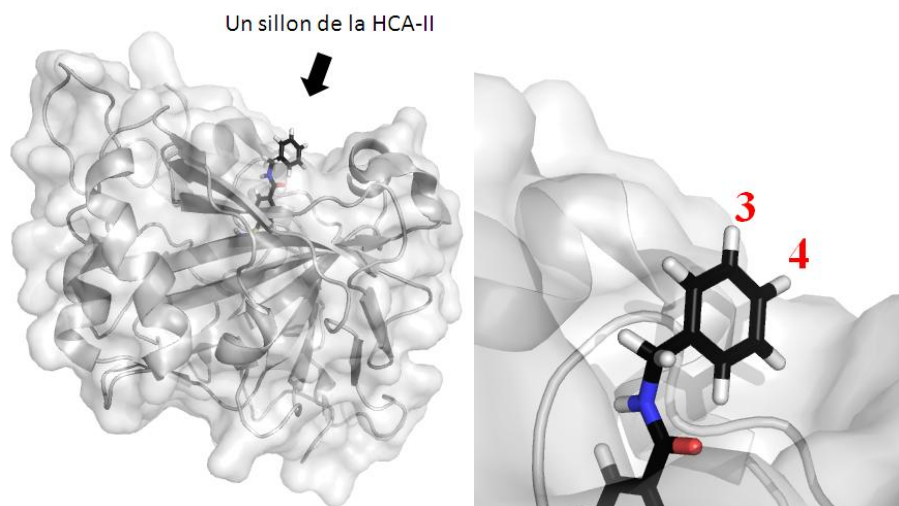
L'inhibiteur SBB (formule en vert), muni d'un espaceur, constitue le système d'ancrage. Le foldamère quant à lui est fonctionnalisé de manière à créer des interactions spécifiques avec la protéine. R représente les chaînes latérales (Thèse de Marine Stupfel, Bordeaux 2010)

La première étape de la stratégie d'ancrage consiste à concevoir un espaceur pertinent entre l'inhibiteur SBB et la première unité quinoline du foldamère à l'aide de la modélisation moléculaire. Il conviendra de vérifier que la fixation de l'espaceur à l'inhibiteur ne modifie pas notablement l'affinité de ce dernier.

### 1.2.2. Conception du système d'ancrage

#### 1.2.2.1. Position de l'espaceur sur l'inhibiteur

La position de l'espaceur sur le ligand SBB choisi est issue des données cristallographiques du complexe SBB/HCA-II (code PDB = 1G4O) (**Figure III.6**). Deux positions, suffisamment dégagées sur le plan stérique pour y placer un espaceur, sont envisageables : l'une en *méta* (3) et l'autre en *para* (4) du benzyle terminal.



**Figure III.6. : Choix de la position de l'espaceur sur l'inhibiteur.**

La structure cristallographique du complexe entre la HCA-II (surface et ruban en gris) et l'inhibiteur SBB (en bâtonnet) (code PDB : 1G4O) met en évidence deux positions envisageables pour l'espaceur sur l'inhibiteur, les positions 3 et 4.

Afin de faciliter les étapes de synthèse chimique, et parce que cette position semble la mieux orientée, la position *méta* (ou position 3) a été retenue comme pouvant être la plus pertinente. Un analogue nommé **1**, de l'inhibiteur SBB, portant un groupement hydroxyle en position *méta* du benzyle a été synthétisé. La structure cristallographique de ce nouveau complexe a été déterminée.

#### 1.2.2.2. Structure du complexe **1**/HCA-II

Des cristaux du complexe **1**/HCA-II ont été obtenus par la technique de la goutte suspendue, en utilisant une solution de protéine (0,3 mM) préincubée 3 jours avec 5 équivalents du composé **1** dans le tampon Tris-HCl (pH 8,5). La solution de complexe (1,2  $\mu$ L) a été ajoutée à 1,2  $\mu$ L de la solution précipitante contenant 2,7 M  $(\text{NH}_4)_2\text{SO}_4$ , 150 mM NaCl, 3 mM  $\text{NaN}_3$ , 2 mM 4-(hydroxymercure)benzoate dans 50 mM Tris-HCl (pH 8,5) à température ambiante. Les premiers cristaux apparaissent après 2-3 jours et atteignent leur taille maximale (0,4 x 0,3 x 0,1 mm) en 3 semaines. Les cristaux sont plongés dans une solution de réservoir complétée par 33 % de glycérol pour assurer leur cryo-protection. Les données de diffraction ont été collectées sur la ligne de lumière ID23-2, à l'ESRF (Grenoble).

La structure a été résolue par remplacement moléculaire grâce à la structure modèle 3KS3. Les données concernant la collecte et l'affinement ont été rassemblées dans le **Tableau III.2.**

1/HCA	
<b>Statistiques de la collecte</b>	
Groupe d'espace	$P2_1$
Paramètre de maille (Å, °)	$a = 42.444, b = 41.559, c = 72.064, \alpha = \gamma = 90, \beta = 104.41$
Résolution (Å)	28.82 - 1.00 (1.050 - 1.000)
$R_{\text{merge}}$	0.098 (0.22)
$I/\sigma(I)$	11.15 (3.21)
Complétude	98.04 (99.03)
Multiplicité	3.49 (3.5)
<b>Statistiques en fin d'affinement</b>	
Résolution (Å)	28.82 - 1.00 (1.026 - 1.000)
No. de réflexions	122692 (9121)
$R_{\text{work}}$	0.1056
$R_{\text{free}}$	0.1226
R.m.s.d	
longueurs de liaison (Å)	0.019
angles de liaison (°)	1.898
Hors Ramachandran	0 %
Ramachandran favorable	96.09 %
Facteur d'agitation thermique moyen (Å <sup>2</sup> )	
protéine	9.26
composé <b>1</b>	12.35
H <sub>2</sub> O	27.15
Nombre de molécules	
H <sub>2</sub> O	439
glycérol	4
pHMB	1
composé <b>1</b>	3

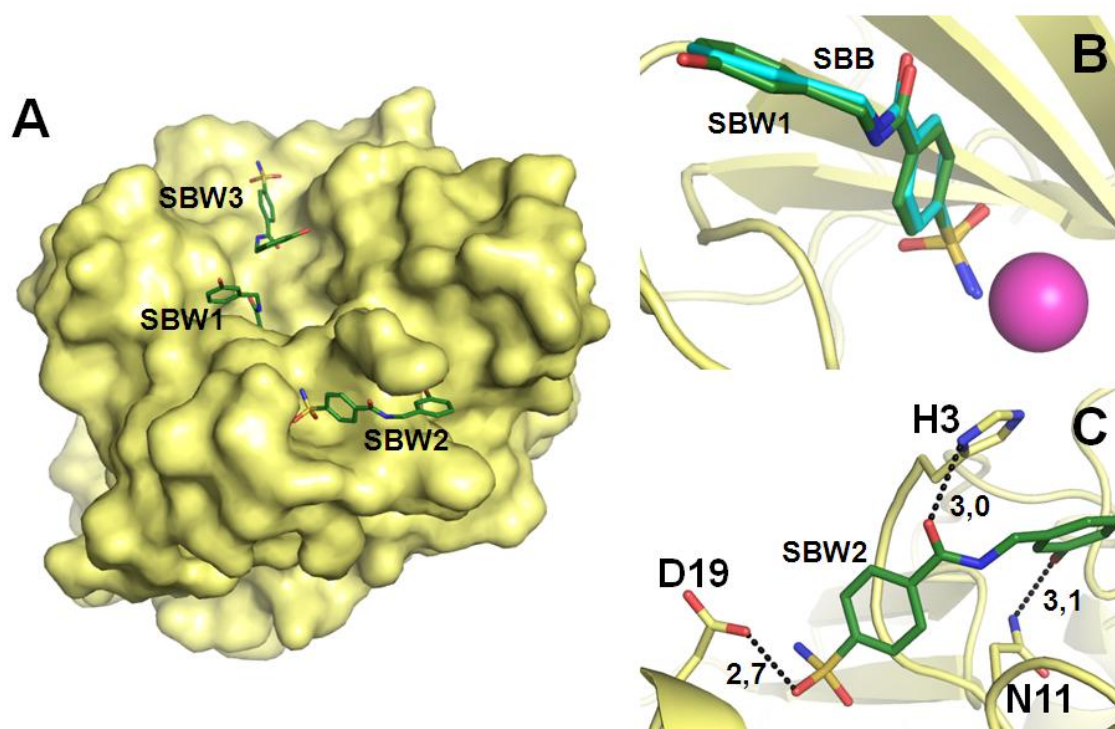
**Tableau III.2. : Statistiques de la collecte de données et de l'affinement.**

Les données entre parenthèses correspondent à la dernière couronne de résolution.

La structure résolue à 1,0 Å de résolution a été déposée à la PDB (code PDB : 4MTY). Compte tenu de la résolution obtenue, le facteur d'agitation thermique de chaque atome a pu être affiné de façon anisotrope. Les cartes de densité ont permis de placer le résidu H3 qui n'est pas visible dans la structure utilisée comme modèle. Il est également important de souligner que la structure comprend trois inhibiteurs **1** (**Figure III.7.A**). Le premier (SBW1) se trouve au site actif, son taux d'occupation est de 1. Sa position est très proche de celle observée pour l'inhibiteur SBB (r.m.s.d = 0,25 Å) (**Figure III.7.B**). Un second composé (SBW2) interagit avec la fonction imidazole de H3 par  $\pi$ -stacking et par liaisons hydrogène avec cette même H3, mais aussi avec D19 et N11 (**Figure III.7.C**). Il a un taux d'occupation de 0,8. Cette seconde molécule stabilise H3 qui peut alors être positionnée dans les cartes de



densité. La troisième molécule **1** (SBW3) est proche de F131, E69 et D72. Elle se trouve exactement dans le sillon 1 (en bleu sur la **Figure III.4.A**). Son taux d'occupation est voisin de 0,3.



**Figure III.7. : Le complexe 1/HCA-II.**

**A :** Trois molécules **1** sont retrouvées dans la structure du complexe. **B :** La superposition de l'inhibiteur modifié (SBW1 en vert) et de SBB (en bleu clair). **C :** Le second inhibiteur **1** (SBW2) interagit avec D19, H3 et N11 principalement. Les longueurs sont données en Å.

Cette structure permet de vérifier que la fonction hydroxyle du composé **1** occupant le site actif s'oriente dans la direction du sillon 2 (en vert sur la **Figure III.4.A**) de la protéine et nous permet, ainsi, de valider la position *méta* du cycle benzyle comme étant la plus judicieuse pour y attacher un espaceur relié au foldamère.

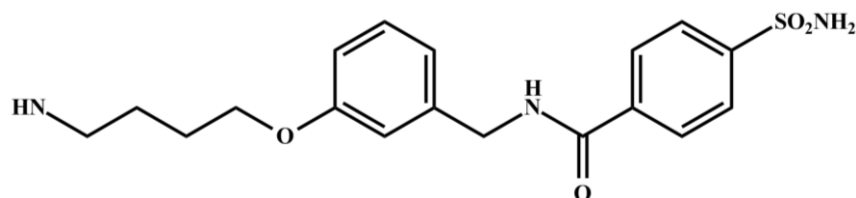
#### 1.2.2.3. Choix de l'espaceur et de la séquence du foldamère cible

Le choix de l'espaceur et la fonctionnalisation du foldamère sont des étapes clés qui doivent être réalisées en tenant compte de divers paramètres structuraux. L'espaceur doit présenter une certaine flexibilité afin de conférer au foldamère une liberté de mouvement lui permettant d'interagir avec la surface de la protéine. Sa longueur doit être définie en fonction de la distance entre le site actif et la surface à reconnaître. Enfin la fonctionnalisation doit être

réalisée de façon à optimiser (dans la limite des monomères pouvant être synthétisés) les interactions avec les résidus présents à la surface de la protéine.

Afin d'étudier la stabilité d'interaction entre les deux partenaires, des études de modélisation ont été réalisées dans le groupe de Michel Laguerre (IECB, Bordeaux) en paramétrant le foldamère avec le champ de force CHARMM General Force Field (CgenFF), et la protéine avec le champ de force CHARMM- 22/CMAP.

Les résultats obtenus montrent qu'un espaceur lié à l'oxygène en position *méta* du composé **1**, et ayant une longueur de 5 liaisons entre l'oxygène de l'inhibiteur et le premier atome du foldamère, permettrait à ce dernier d'interagir convenablement avec la protéine dans le sillon souhaité. L'inhibiteur modifié sur lequel est fixé l'espaceur forme le composé **2** (**Figure III.8**).



**Figure III.8 : Formule de la molécule 2, également nommée Inh.**

Le système d'ancrage et l'espaceur ayant été choisis, les calculs de dynamique moléculaire ont été effectués sans qu'aucune contrainte ne soit appliquée entre le foldamère et la protéine. Ces simulations ont permis d'optimiser la longueur et la fonctionnalisation du foldamère en considérant un volume de simulation uniquement constitué de molécules d'eau. Les résultats obtenus sont présentés **Figure III.9**. Un dodécamère constitue le foldamère cible, optimisant les interactions de surface avec l'anhydrase carbonique. Quatre interactions spécifiques, sont observées entre les unités 1, 3, 5, 6, 8, 10 et 11 du dodécamère de quinoline et les résidus H1, R58, E69, Q92, D130, F131, K133 et K170 de la surface de la protéine. Ces interactions sont de type hydrophobe (F131 avec Q<sub>3</sub><sup>Leu</sup>) ou bien impliquent des liaisons hydrogène (K170 avec Q<sub>1</sub><sup>Asp</sup>). Le dodécamère cible est représenté **Figure III.10**.



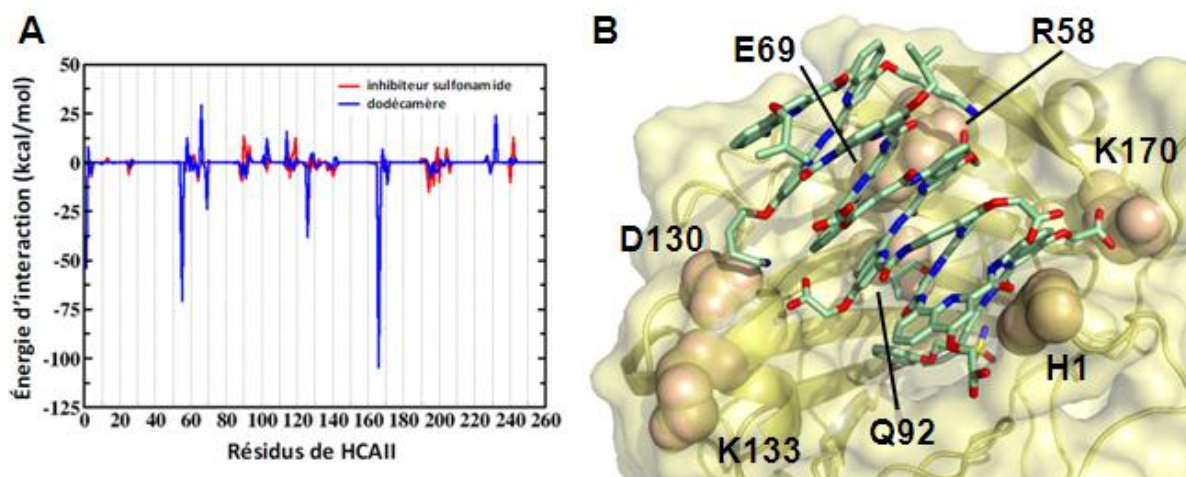


Figure III.9. : Résultats de la modélisation moléculaire ayant permis la conception du foldamère cible.

A : Energie d'interaction entre la HCA-II et le dodécamère cible. B : Résultat de la modélisation montrant le foldamère (bâtonnets) interagissant avec les résidus clés (sphères jaunes) de la surface de la HCA-II (en jaune pâle).

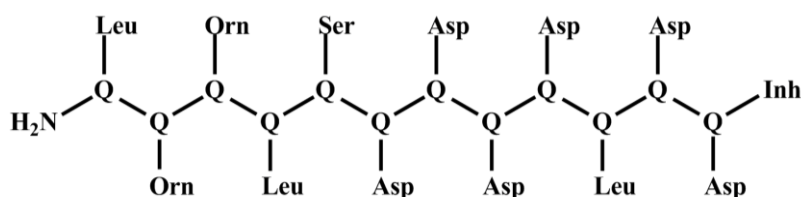


Figure III.10. : Formule du dodécamère de quinoline cible.

Q représente le monomère de quinoline. Ici le foldamère est ancré à l'espaceur (molécule 2) par son extrémité C-ter via une liaison amide.

#### 1.2.2.4. Effet de l'addition de l'espaceur sur l'inhibiteur modifié.

Afin de s'assurer que l'espaceur ne perturbe pas l'affinité de l'inhibiteur modifié **1** pour le site actif de l'enzyme, des expériences de SPR ont été menées (**Figure III.11.**). Nous avons pu caractériser cinétiquement l'association et la dissociation de l'inhibiteur en présence ou non de l'espaceur. Le composé avec l'espaceur (**2**) a été synthétisé par l'équipe de chimistes d'Ivan Huc. Les conditions expérimentales des expériences de SPR sont décrites dans la partie IV. Matériels et méthodes, paragraphe 3.3.

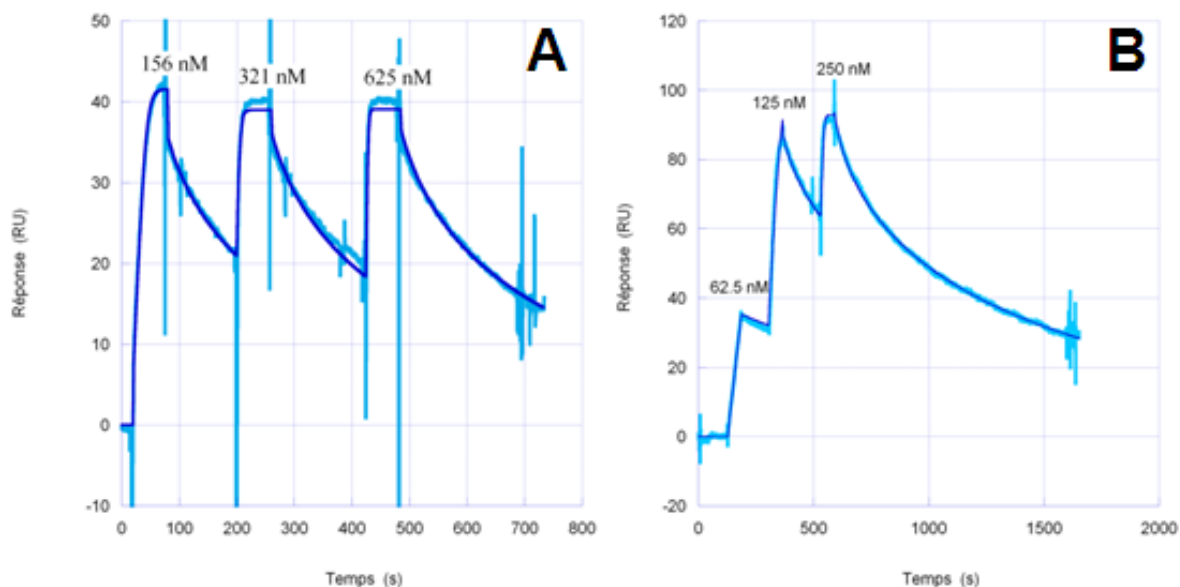


Figure III.11. : Résultats des expériences de SPR : A : Composé 1, B : Composé 2

Les concentrations en composé injecté sont notées au dessus de chaque pic. Les données brutes sont en bleu clair. Le paramétrage obtenu en utilisant un modèle de liaison 1:1 est représenté en bleu foncé.

L'analyse des sensorgrammes de titration des composés **1** et **2**, montre que l'ajout de l'espaceur à l'inhibiteur modifié, n'a pas d'effet sensible : les constantes d'association ( $k_{on}$ ), de dissociation ( $k_{off}$ ) et d'affinité ( $K_D$ ) des composés **1** et **2** pour la HCA-II sont rassemblées dans le **Tableau III.3**.

	<b>1</b>	<b>2</b>
$k_{on} (M^{-1}.s^{-1})$	$7,1 \cdot 10^5$	$10,3 \cdot 10^5$
$k_{off} (s^{-1})$	$5,2 \cdot 10^{-3}$	$5,6 \cdot 10^{-3}$
$K_D (nM)$	7,3	5,4

Tableau III.3. : Constantes cinétiques des interactions entre les composés **1** ou **2** et la HCA-II.

Ces résultats valident définitivement notre stratégie d'utiliser un inhibiteur de l'enzyme comme point d'ancrage pour maintenir le foldamère à proximité de la surface cible de la protéine et ainsi favoriser des interactions entre les deux partenaires.

## 2. Les foldamères de première génération

Rappelons que nous qualifions de foldamères de première génération les foldamères synthétisés en solution. Une telle méthode de synthèse impose que l'inhibiteur et l'espaceur associé soient greffés en partie C-terminale du foldamère par une liaison amide.

### 2.1. Bibliothèque

Quatre foldamères (un trimère, un pentamère, et deux heptamères) dont les séquences sont mentionnées dans le **Tableau III.4.** ont été étudiés. Les séquences des composés **3**, **4** et **6** correspondent à la séquence tronquée du dodécamère cible cité dans la partie précédente. Les chimistes de l'équipe d'Ivan Huc ont pensé introduire une petite variabilité au niveau de l'avant dernière chaîne latérale, accroissant potentiellement les chances de voir une interaction. Nous avons travaillé avec des solutions stock de ces composés à 10 mM dans le DMSO car leur solubilité dans l'eau n'est pas suffisante pour nos besoins.

n° du composé	Séquence	Masse Molaire (Da)
<b>3</b>	Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Inh	1150,19
<b>4</b>	Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Inh	1626,53
<b>5</b>	Q <sup>Asp</sup> -Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Inh	2083,92
<b>6</b>	Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup> -Inh	2084,95

**Tableau III.4. : Séquences et masses molaires des 4 composés de la première génération.**

### 2.2. Evaluation des interactions foldamère/protéine

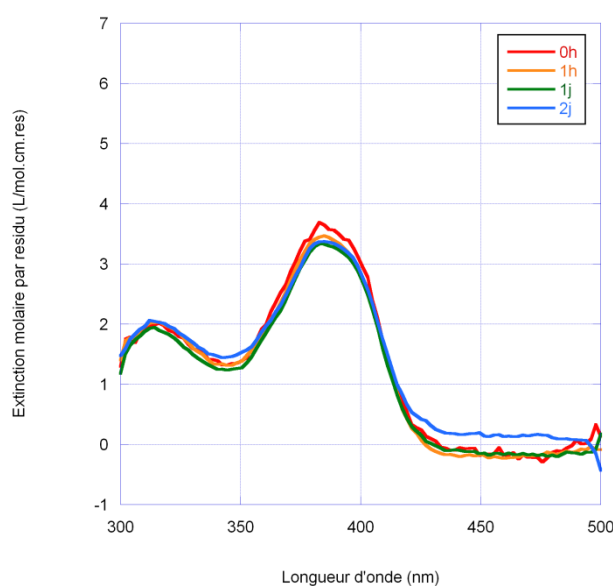
#### 2.2.1. Principe du Dichroïsme Circulaire Induit (ICD)

Rappelons que l'interaction entre un foldamère et une protéine peut être stéréosélective et peut modifier l'équilibre hélice droite / hélice gauche observé pour un foldamère en solution. Dans ce cas, sera observé un signal de dichroïsme circulaire induit

(voir partie IV. Matériels et méthodes, paragraphe 3.1.). Ce signal traduit la prépondérance d'une hélicité donnée, induite par l'interaction entre partenaires. Dans le cas de foldamères de quinolines, le pic d'absorption attendu se situe vers 390 nm (Kendhale *et al.* 2011).

### 2.2.2. Induction de chiralité du trimère **3** par la HCA-II

Des expériences de dichroïsme circulaire induit ont été conduites afin de mesurer l'induction de chiralité de la protéine sur le trimère **3** qui, compte tenu de sa longueur, se structure sous la forme d'un seul tour d'hélice. Afin d'être sûr que l'inversion d'hélicité ait pu se produire, nous avons mesuré le même spectre après différents temps d'incubation dans la gamme de longueurs d'onde : 300 nm – 500 nm. L'incubation s'effectue à 4°C de façon à maintenir l'intégrité de la protéine. Dans tous les cas, une cuve de 2 mm d'épaisseur est utilisée et le rapport foldamère : protéine est toujours maintenu égal à 1. Au spectre du complexe obtenu est systématiquement soustrait le spectre de la protéine seule, de manière à n'observer que le signal correspondant au foldamère. Les résultats sont présentés dans la **Figure III.12.**



**Figure III.12. : Spectres ICD du complexe 3/HCA-II obtenus à 25 °C.**

Les mesures ont été réalisées dans un tampon  $\text{NaH}_2\text{PO}_4$  50 mM, pH 7,5. Les temps d'incubation sont mentionnés en heures et en jours en encart.

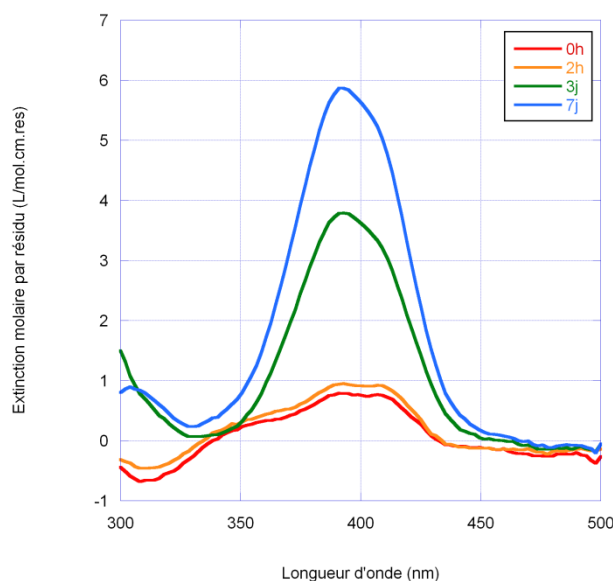
A  $t=0$ , c'est-à-dire immédiatement après que le mélange soit effectué, un pic d'absorption à 390 nm est observé. Son intensité ne varie pas en fonction du temps

d'incubation. Ce signal correspond à une absorbance positive et traduit, d'après les résultats publiés par Kendhale *et al.* (2011), la prépondérance d'une hélicité droite.

Ces données nous permettent donc d'affirmer que le trimère **3** interagit avec la surface de la protéine, sans qu'il nous soit toutefois possible de préciser l'affinité d'interaction. Afin de vérifier si un foldamère de séquence plus longue pouvait également interagir avec la protéine, de nouveaux spectres de dichroïsme circulaire induit ont été enregistrés en utilisant les autres composés du **Tableau III.4**.

### 2.2.3. Induction de chiralité du pentamère **4** par la HCA-II

Les conditions expérimentales précédentes ont été conservées (enregistrement des spectres à 25 °C, même traitement du signal, même cuve et même rapport stœchiométrique entre foldamère et protéine). Les spectres obtenus sont présentés dans la **Figure III.13**. Notons qu'ils ont cependant été enregistrés après des temps d'incubation plus longs.



**Figure III.13. : Spectres ICD du complexe 4/HCA-II obtenus à 25 °C.**

**Les mesures ont été réalisées dans un tampon NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 7,5. Les différents temps d'incubation sont mentionnés en heures et en jours en encart.**

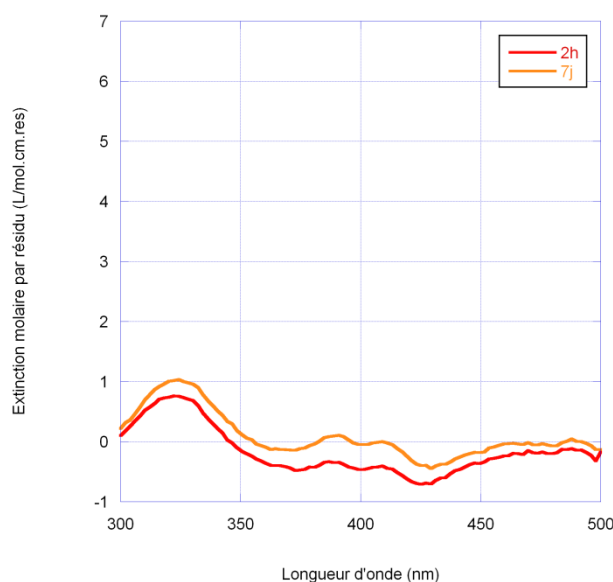
Nous avons émis l'hypothèse que plus la longueur du foldamère était conséquente (2 tours d'hélice dans ce cas au lieu d'un seul précédemment), plus le temps d'incubation nécessaire pour adopter une hélicité préférentielle gauche ou droite serait grand. Le maximum de signal est effectivement atteint après 7 jours d'incubation. Ceci confirme notre hypothèse.

Le signal d'environ +6 L/mol.cm.res est beaucoup plus important que celui du trimère **3** (environ +3,5 L/mol.cm.res).

Ce résultat tend à montrer qu'une élongation de la séquence du foldamère peut favoriser une plus forte interaction avec la surface de la protéine dès lors que la fonctionnalisation de ce dernier reste favorable.

#### 2.2.4. Induction de chiralité de l'heptamère **5** par la HCA-II

Les spectres de dichroïsme circulaire induit de l'heptamère **5** (**Figure III.14.**) ont été mesurés dans les mêmes conditions que précédemment.



**Figure III.14. : Spectres ICD du complexe **5**/HCA-II enregistrés à 25°C.**

**Les mesures ont été réalisées dans un tampon NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 7,5. Les différents temps d'incubation sont mentionnés en heures et en jours encart.**

Aucun pic d'absorption n'est visible dans le domaine 300 – 500 nm. A la vue de ces résultats, nous supposons que le temps d'incubation n'était pas suffisant pour induire un changement d'hélicité ou que les fonctions portées par les nouvelles unités quinolines s'opposaient à toute interaction. L'équipe d'Ivan. Huc, travaillant sur la cinétique d'inversion d'hélicité de foldamères de quinoline en milieu aqueux, a montré que le temps nécessaire au changement d'hélicité était fonction de la longueur du foldamère (Delsuc *et al.* 2008) et qu'à partir de huit unités quinolines, le foldamère présente une stabilité si importante dans un solvant protique, que l'interconversion est virtuellement impossible (donnée non publiée). Il

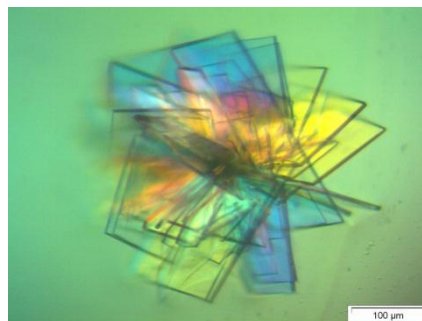
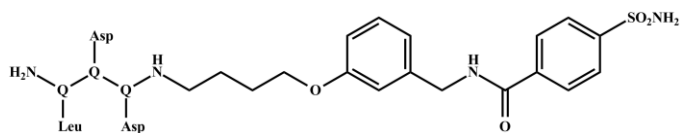
devient donc difficile, par dichroïsme circulaire, de conclure quant à une possible interaction entre la protéine et l'heptamère étudié.

## 2.3. Tentatives de cristallisation

Au vue des résultats obtenus en dichroïsme circulaire induit, nous avons tenté de cristalliser les trois complexes **3**/HCA-II, **4**/HCA-II et **5**/HCA-II, afin d'obtenir des informations structurales sur les interactions foldamère - protéine.

### 2.3.1. Cristallisation du complexe **3**/HCA-II

De nombreuses expériences de criblage (Crystal Screen, SaltRx, JCSG+, JBScreen 6) ont été conduites en faisant varier le nombre d'équivalents du composé **3** par rapport à la protéine (1,2 à 10 équivalents), ainsi que la température (4, 13, 20 et 30 °C). Des plaquettes cristallines apparaissent au bout d'une semaine dans la condition C5 du *screen* JBScreen 6 à 20 °C (2,2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>) (**Figure III.15.**). Après optimisation par la méthode de diffusion de vapeur en gouttes suspendues, des cristaux en forme de plaquette (100 x 200 µm) ont été obtenus à 20 °C. Pour cela, la protéine (0,3 mM dans 50 mM Tris-HCl, pH 8,0) a été incubée avec **3** (2,4 eq) pendant deux jours à 4 °C, puis 1,2 µL de cette solution ont été ajoutés à 1,2 µL de la solution de réservoir contenant 310 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 3 mM NaN<sub>3</sub> dans 25 mM AcNa (pH 4,8). Les cristaux obtenus ont été cryo protégés par trempage dans une solution de réservoir complémentée par 33 % de glycérol. Ils ont été testés sur la ligne de lumière Proxima à Soleil (Gif sur Yvette). Aucun d'entre eux ne diffractait à une résolution meilleure que 4 Å. La stabilité limitée de ces cristaux, alliée à une anisotropie importante du pouvoir diffractant, rendait en outre difficile l'enregistrement d'un jeu complet d'intensités diffractées.

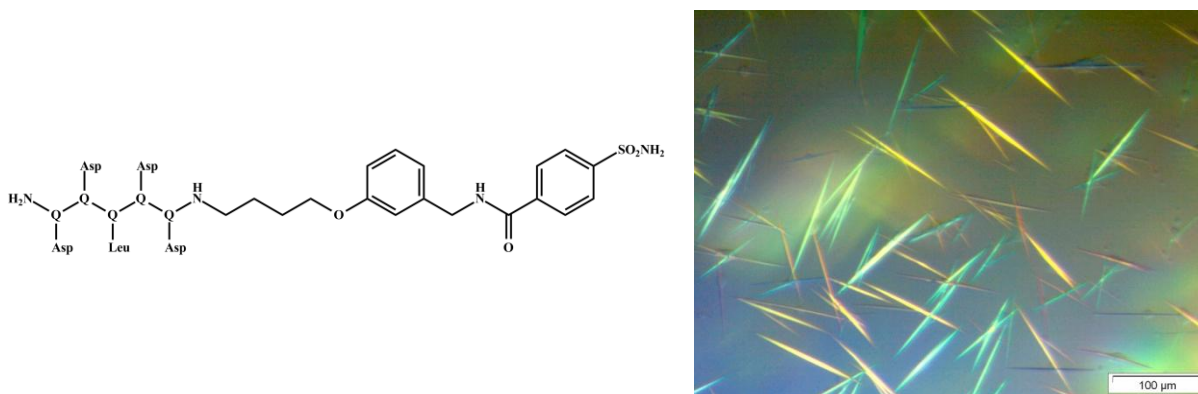


**Figure III.15.** Formule simplifiée du composé **3** et cristaux du complexe **3**/HCA-II.

Photo en lumière polarisée d'un « amas » de plaquettes cristallines obtenues après optimisation.

### 2.3.2. Cristallisation du complexe 4/HCA-II

Des tentatives de cocrystallisation (Crystal Screen, SaltRx, JCSG+, JBScreen 6), ont été menées en faisant varier le nombre d'équivalents du composé **4** par rapport à la protéine (1,2 à 2,4 équivalents), ainsi que la température (4 et 20 °C). Seules des aiguilles cristallines sont apparues (3 semaines) dans la condition B5 de JBScreen 6 (1,6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 % PEG 1000 dans 100 mM HEPES (pH 7,5) à 20°C avec une concentration en protéine dans la goutte de 0,3 mM (dans 50 mM Tris-HCl, pH 8,0) et un rapport foldamère : protéine de 2,4 (**Figure III.16**). Toutes les tentatives d'optimisation pour obtenir des cristaux de plus grandes dimensions sont restées vaines.



**Figure III.16. : Formule simplifiée du composé 4 et cristaux du complexe 4/HCA-II.**

**Photo en lumière polarisée des aiguilles cristallines obtenues.**

### 2.3.3. Cristallisation des heptamères 5 et 6 avec la HCA-II

Le criblage des conditions de cristallogénèse des complexe **5**/HCA-II et **6**/HCA-II résumées dans le **Tableau III.5**, a permis de dégager de nombreuses conditions où des aiguilles et des oursins sont apparus. La concentration de la solution de protéine utilisée était de 0,3 mM dans le tampon usuel et le rapport foldamère : protéine de 2,4. Les tentatives d'optimisation de concentration en sel, de pH ou d'agents précipitants ont été vaines.



Composé utilisé	Screen	Position	Condition	Forme cristalline
6	Classics II	49/E1	0,2 M $\text{CaCl}_2$ , 45 % MPD dans 100 mM Bis-Tris, pH 6,5	oursins
5	ComPas	62/F2	0,2 M $\text{MgCl}_2$ , 30 % isoPropanol dans 100 mM HEPES, pH 7,5	aiguilles
6	ComPas	58/E10	0,2 M $\text{MgCl}_2$ , 20 % isoPropanol dans 100 mM NaAc, pH 4,6	aiguilles
6	ComPas	60/E12	0,2 M $\text{MgCl}_2$ , 25 % isoPropanol dans 100 mM HEPES, pH 7,5	aiguilles
6	ComPas	62/F2	0,2 M $\text{MgCl}_2$ , 30 % isoPropanol dans 100 mM HEPES, pH 7,5	aiguilles
6	ComPas	64/F4	0,1 M $\text{MgCl}_2$ , 25 % tertButanol dans 100 mM Tris-HCl, pH 8,5	oursins
5	JCSG Core II	85/H1	0,2 M $\text{Ca}(\text{Ac})_2$ , 30 % PEG 400 dans 100 mM NaAc, pH 4,5	oursins
6	JCSG core II	77/G5	0,2 M $\text{CaCl}_2$ , 20 % PEG 3 350	nodules
5	Mb Class	38/D2	0,2 M $\text{CaCl}_2$ , 42,6 % PEG 400 dans 100 mM HEPES, pH 7,5	sels
6	Mb Class	3/A3	0,2 M $\text{CaCl}_2$ , 38,2 % isoPropanol, 15,9 % Glycérol dans 100 mM NaAc, pH 4,6	aiguilles
6	Mb Class	38/D2	0,2 M $\text{CaCl}_2$ , 42,6 % PEG 400 dans 100 mM HEPES, pH 7,5	sels
6	MPD	45/D9	0,2 M 3LiCit, 40 % MPD	nodules
5	PEG II	36/C12	0,2 M $\text{CaCl}_2$ , 20 % PEG 4 000 dans 100 mM Tris, pH 8,5	oursins
6	pH Clear	72/F12	3,2 M $(\text{NH}_4)_2\text{SO}_4$ , dans 100 mM Bicine, pH 9,0	nodules, aiguilles

Tableau III.5. : Conditions de *screen* ayant fait apparaître des cristallites.

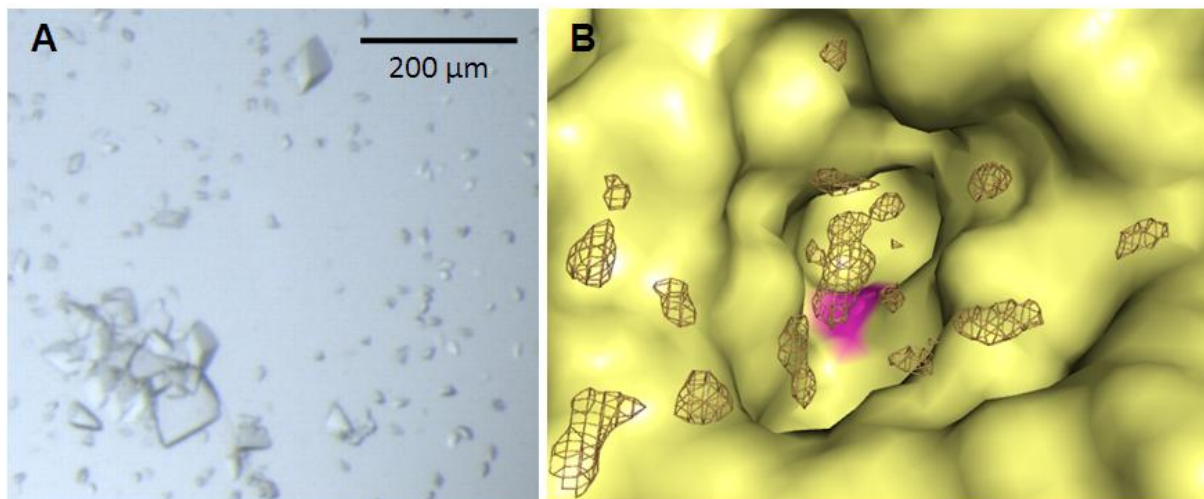
Devant ces pistes, nous avons donc tenté des expériences d'ensemencement avec des germes provenant de cristaux de HCA-II native. Grâce à cette technique, nous avons pu obtenir des cristaux diffractant du complexe 5/HCA-II. La méthode de diffusion de vapeur en goutte suspendue a été utilisée. Une condition proche des conditions classiques de cristallisation de la HCA-II a été choisie car aucun précipité, nodule ou séparation de phase n'était observé avec ce complexe. Le mélange de la solution de complexe (1,2  $\mu\text{L}$  de HCA-II

à 0,3 mM dans son tampon habituel avec 2,4 équivalents de **5**) avec la solution de réservoir (1,2  $\mu$ L à 150 mM NaCl, 2,7 M  $(\text{NH}_4)_2\text{SO}_4$ , 2 mM pHMB dans 100 mM Tris (pH 8,5)) a été équilibré pendant une journée entière. Un poil de chat ayant été trempé dans une goutte de cristallisation contenant des cristaux de HCA-II native est utilisé pour réaliser un *streak*, ou trait dans la goutte équilibrée. Les cristaux apparaissent un jour après l'ensemencement. Leur cryo-protection s'est effectué par trempage dans une solution de réservoir complétée par 33 % de glycérol. La collecte du jeu de données a été faite sur la ligne de lumière I04 à Diamond (Oxford). La résolution de la structure a été faite par remplacement moléculaire avec comme modèle le fichier PDB de code 3KS3. Les données concernant la collecte et l'affinement ont été rassemblées dans le **Tableau III.6**.

5/HCA	
Statistiques de la collecte	
Groupe d'espace	$P2_1$
Paramètre de maille (Å, °)	$a = 42,160$ , $b = 41,640$ , $c = 72,410$ , $\alpha = \gamma = 90$ , $\beta = 104,20$
Résolution (Å)	41,64 - 1,23 (1,260- 1,230)
$R_{\text{merge}}$	0,054 (0,516)
$I/\sigma(I)$	11,3 (2,3)
Complétude	99,1 (99,8)
Multiplicité	3,6 (3,6)
Statistiques en fin d'affinement	
Résolution (Å)	41,64 - 1,23 (1.262 - 1.230)
No. de réflexions	66 651 (4 807)
$R_{\text{work}}$	18,76 (26,9)
$R_{\text{free}}$	21,32 (30,4)
R.m.s.d	
longueurs de liaison (Å)	0,029
angles de liaison (°)	2,479
Hors Ramachandran	0 %
Ramachandran favorable	94,42 %
Facteur d'agitation thermique moyen (Å <sup>2</sup> )	
protéine	14,88
H <sub>2</sub> O	23,22
Nombre de molécules	
H <sub>2</sub> O	162
glycérol	1
pHMB	1
composé <b>5</b>	0

**Tableau III.6. : Statistiques de la collecte de données et de l'affinement.**

Les données entre parenthèses correspondent à la dernière couronne de résolution.



**Figure III.17. : Cristaux et site actif du complexe 5/HCA-II.**

**A :** Image de la goutte de cristallisation une semaine après l'ensemencement. **B :** Surface de la HCA-II dans le complexe 5/HCA-II, et la carte de densité électronique mFo-DFc contournée à  $3\sigma$  au voisinage du site actif. La position de l'atome de zinc est localisée à la surface de la protéine par sa couleur violette.

La carte mFo-DFc contournée à  $3\zeta$  au voisinage du site actif (**Figure III.17.**) fait uniquement apparaître des densités résiduelles qui ne peuvent être corrélées à la présence de composé. L'inhibiteur et *a fortiori* l'espaceur et le foldamère ne sont pas visibles. Il est possible que l'affinité de l'inhibiteur modifié ait été diminuée par la présence de l'heptamère. Des expériences de biophysique (SPR, ITC, RMN, ...) sont toutefois nécessaires pour confirmer cette hypothèse.

Il aurait été nécessaire de poursuivre les efforts engagés dans la cristallogénèse des complexes foldamère/protéine précédents et de compléter les informations obtenues en utilisant de nouvelles techniques d'étude. Ceci aurait nécessité de disposer d'une quantité plus importante de foldamères produits. La mise au point d'une nouvelle méthode de synthèse, plus performante tant en quantité de composés produits qu'en rapidité de mise en œuvre, a conduit les chimistes à abandonner les productions précédentes trop coûteuses et peu productives. Ceci nous a inévitablement amenés à reprendre les études précédentes sur de nouveaux composés.

## 3. Les foldamères de seconde génération

Ces nouveaux foldamères sont obtenus par synthèse sur support solide. Cette technique permet de conduire des synthèses de façon plus rapide, plus efficace en termes de quantité de produits obtenus. Elle permet en outre une plus grande flexibilité de fonctionnalisation des séquences synthétisées. Elle s'avère donc beaucoup plus adaptée à l'application attendue, à savoir, la production de foldamères spécifiquement « décorés » pour reconnaître une surface donnée. Cette technique implique toutefois un ancrage des foldamères à l'espaceur non plus par une liaison amide à leur partie C-terminale mais par une liaison urée à leur extrémité N-terminale.

### 3.1. Bibliothèque

Les foldamères de « seconde génération » que nous avons étudiés sont au nombre de 10 : deux trimères, quatre tétramères, et quatre pentamères. Ils ont été synthétisés par l'équipe d'Ivan Huc. Leurs séquences sont présentées dans le **Tableau III.7.**

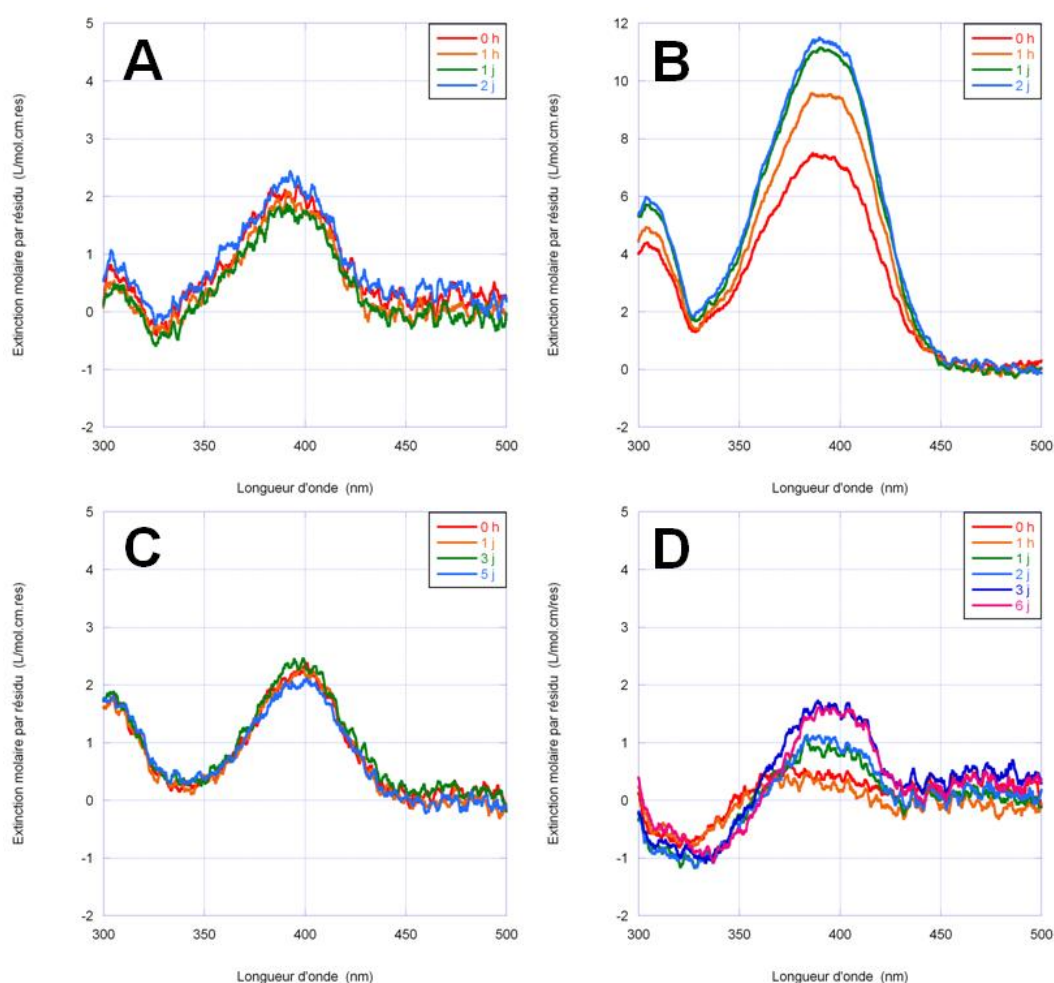
n° du composé	Séquence	Masse Molaire (Da)
<b>7</b>	Inh-Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup>	1151,15
<b>8</b>	Inh-Q <sup>Leu</sup> -Q <sup>Hyd</sup> -Q <sup>Orn</sup> -Q <sup>Asp</sup>	1337,37
<b>9</b>	Inh-Q <sup>Hyd</sup> -Q <sup>Hyd</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup>	1282,21
<b>10</b>	Inh-Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Asp</sup>	1640,55
<b>11</b>	Inh-Q <sup>Leu</sup> -Q <sup>Orn</sup> -Q <sup>Orn</sup> -Q <sup>Ala</sup>	1350,46
<b>12</b>	Inh-Q <sup>Asp</sup> -Q <sup>Asp</sup> -Q <sup>Leu</sup>	1152,15
<b>13</b>	Inh-Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Ala</sup> -Q <sup>Phe</sup> -Q <sup>Orn</sup>	1627,63
<b>14</b>	Inh-Q <sup>Leu</sup> -Q <sup>Tyr</sup> -Q <sup>Asp</sup> -Q <sup>Ala</sup> -Q <sup>Orn</sup>	1659,75
<b>15</b>	Inh-Q <sup>Leu</sup> -Q <sup>Asp</sup> -Q <sup>Tyr</sup> -Q <sup>Ala</sup> -Q <sup>Orn</sup>	1659,75
<b>16</b>	Inh-Q <sup>Ala</sup> -Q <sup>Asp</sup> -Q <sup>Phe</sup> -Q <sup>Orn</sup>	1385,41

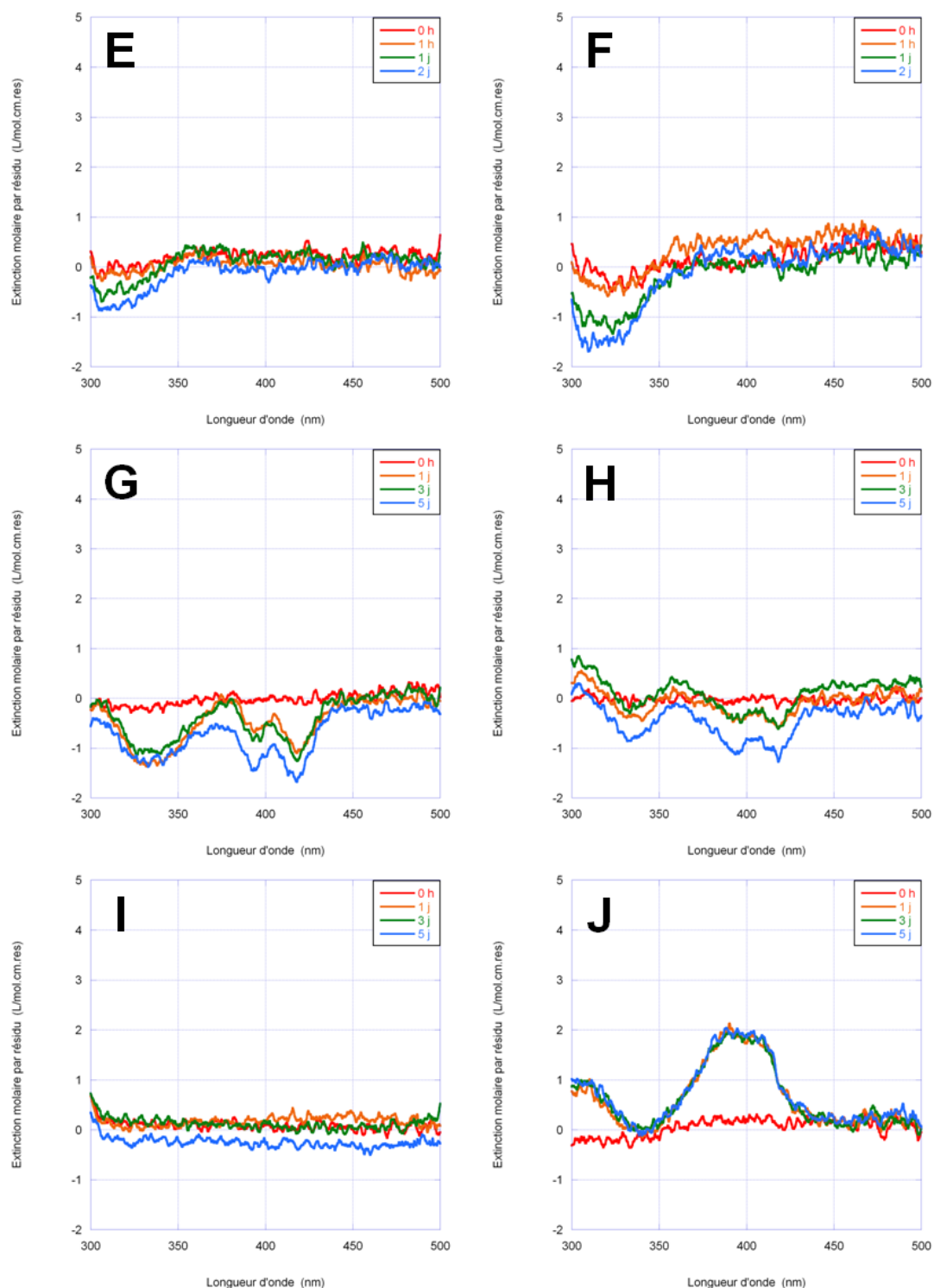
**Tableau III.7. : Séquences et masses molaires des 10 composés de seconde génération.**

## 3.2. Identification d'un candidat potentiel par analyses biophysiques

### 3.2.1. Dichroïsme circulaire induit

De la même manière qu'avec les composés de première génération, nous avons souhaité savoir si l'un ou plusieurs de ces nouveaux composés présentait une hélicité préférentielle en présence de la protéine. Nous avons donc réalisé des études de dichroïsme circulaire semblables à celles précédemment menées. Les conditions expérimentales sont en tout point conservées (temps d'incubation à 4 °C variable, enregistrement des spectres à température ambiante, rapport des concentrations 1:1, ...). Les spectres d'ICD, présentés **Figure III.18**, sont des spectres différence traduisant la seule contribution du foldamère (spectre de la protéine seule retranché au spectre du complexe).





**Figure III.18. : A à J : Spectres ICD des composés 7 à 16 complexés à la HCA-II respectivement. Les différents temps d'incubation sont indiqués sur chaque spectre en heures et en jours en encart.**

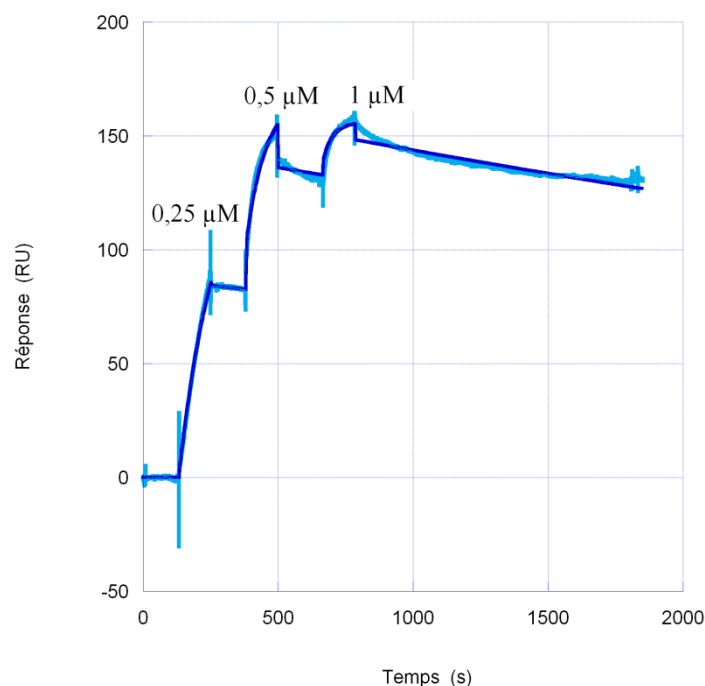
Tous ces spectres, à l'exception d'un seul, montrent un signal de très faible intensité traduisant une influence très limitée de la protéine sur l'hélicité des foldamères. Rappelons

que l'absence de tout signal ne signifie pas l'absence d'interactions entre les deux partenaires. Il traduit seulement le fait que ces interactions, si elles existent, ne sont pas stéréosélectives, c'est-à-dire qu'elles n'impliquent pas la présence d'un conformère donné.

Le composé **8** quant à lui affiche un signal important à 390 nm. Son amplitude est proche de +12 L/mol.cm.res, valeur sensiblement 2 fois plus importante que celle du composé **4** (+6 L/mol.cm.res) étudié précédemment (voir **Figure III.13.**). Ce signal positif indiquerait que le foldamère adopte préférentiellement une hélicité droite.

### 3.2.2. Analyses par Résonance Plasmonique de Surface (SPR)

Afin de savoir si la présence du foldamère  $Q^{\text{Leu}}-Q^{\text{Hyd}}-Q^{\text{Om}}-Q^{\text{Asp}}$  (composé **8**) à l'extrémité de l'inhibiteur améliore ou altère les constantes cinétiques d'association et de dissociation avec la protéine, nous avons réalisé des expériences de SPR (**Figure III.19.**). Les conditions expérimentales sont celles mentionnées dans la partie IV. Matériels et méthodes, paragraphe 3.3.



**Figure III.19. : Sensorgramme de titration du composé 8.**

Les concentrations en composé injecté sont notées au dessus de chaque pic. Les données brutes sont en bleu clair. Le paramétrage obtenu en utilisant un modèle de liaison 1:1 est représenté en bleu foncé.

La constante d'affinité calculée ( $K_D = 4,4$  nM) est identique à celles mesurées pour les composés **1** (inhibiteur seul) et **2** (inhibiteur portant l'espaceur). Aussi, l'ancrage du foldamère ne modifie-t-il pas de façon significative l'affinité de l'inhibiteur pour la protéine. Ces résultats sont reportés dans le **Tableau III.8**. Il est cependant très important de noter que les constantes d'association et de dissociation du composé **8** sont environ 20 fois inférieures à celles mesurées avec les composés **1** et **2**. Le ralentissement de la vitesse d'association pourrait être expliqué par la présence du foldamère sur l'inhibiteur. La vitesse de dissociation est elle aussi ralentie. Des interactions entre le foldamère et la surface de la HCA-II pourraient en être responsables, comme semblent l'indiquer les résultats de CD.

	<b>1</b>	<b>2</b>	<b>8</b>
$k_{on}$ ( $M^{-1}.s^{-1}$ )	$7,1 \cdot 10^5$	$10,3 \cdot 10^5$	$4,5 \cdot 10^4$
$k_{off}$ ( $s^{-1}$ )	$5,2 \cdot 10^{-3}$	$5,6 \cdot 10^{-3}$	$2,0 \cdot 10^{-4}$
$K_D$ (nM)	7,3	5,4	4,4

**Tableau III.8. : Constantes d'association, de dissociation, et d'affinité des composés 1, 2 et 8 pour la HCA-II mesurées en SPR.**

Pour obtenir des informations complémentaires sur les possibles interactions entre **8** et HCA-II, nous avons souhaité tirer profit de la supériorité de la méthode de synthèse sur support solide permettant entre autre une fonctionnalisation aisément modulable du composé synthétisé. C'est ainsi que différents foldamères ont été produits, chacun d'eux ne se différenciant du composé **8** que par suppression d'une des chaînes latérales présentes le long du foldamère.

### 3.2.3. *Evaluation du rôle joué par les chaînes latérales*

#### 3.2.3.1. Principe de la méthode utilisée

La méthode retenue est calquée sur celle dite de l'« *Ala scanning* » qui, par mutation d'un résidu donné en alanine, permet de préciser l'implication de ce résidu dans la stabilité ou la fonction d'une protéine donnée (Edelheit 2011). Dans le cas de protéines, le choix de l'alanine est dicté par le fait que ce résidu présente différents atouts (Wells 1991). Sa chaîne latérale (un simple groupement méthyle) est de faible encombrement stérique, ne porte pas de charge, se retrouve dans tout type de structure secondaire et n'engendre pas de changement de repliement comme pourrait le permettre une glycine par exemple.



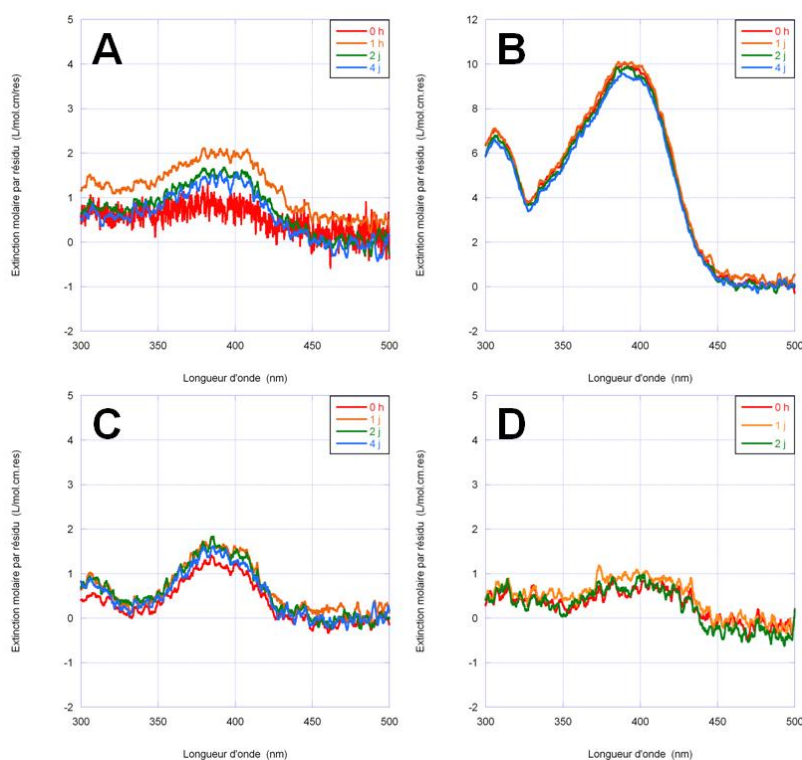
Dans notre cas, nous avons choisi de muter les chaînes latérales par un simple atome d'hydrogène (chaîne latérale de la glycine), puisque la structure secondaire en hélice est imposée par le squelette quinoline et non par la fonction portée par chaque sous-unité. Les composés ainsi obtenus sont mentionnés dans le **tableau III.9**.

n° du composé	Séquence	Masse Molaire (Da)
<b>17</b>	Inh-Q <sup>Leu</sup> -Q <sup>Hyd</sup> -Q <sup>Orn</sup> -Q <sup>Gly</sup>	1263,34
<b>18</b>	Inh-Q <sup>Leu</sup> -Q <sup>Hyd</sup> -Q <sup>Gly</sup> -Q <sup>Asp</sup>	1264,28
<b>19</b>	Inh-Q <sup>Leu</sup> -Q <sup>Gly</sup> -Q <sup>Orn</sup> -Q <sup>Asp</sup>	1321,37
<b>20</b>	Inh-Q <sup>Gly</sup> -Q <sup>Hyd</sup> -Q <sup>Orn</sup> -Q <sup>Asp</sup>	1265,27

**Tableau III.9. : Séquences et masses molaires des 4 composés utilisés.**

### 3.2.3.2. : Impacts des mutations réalisées

L'idée poursuivie est de déterminer si la fonction portée par l'une des sous-unités quinoline est déterminante dans l'interaction avec la protéine. Si tel est le cas, l'analyse des spectres d'IDC devrait nous fournir des éléments de réponse.



**Figure III.20. : Spectres ICD des composés 17 à 20 (respectivement de A à D) en présence de la HCA-II. Les différents temps d'incubation sont donnés en encart.**

Seul le spectre du composé **18** présente toujours un pic à 390 nm (**Figure III.20.**). Toutefois l'intensité de ce pic est légèrement moins intense (+10 L/mol.cm.res) que celle (+12 L/mol.cm.res) du pic observé pour le composé **8**. Ceci tendrait à montrer que la fonction ornithine supprimée dans ce composé ne serait pas essentielle dans l'interaction responsable de l'induction de chiralité du foldamère. Au contraire les autres spectres montrent une chute significative du signal d'ICD, ce qui sous-entend que les chaînes leucine, sérine et aspartate sont d'une manière ou d'une autre impliquées dans les interactions responsables de l'induction de chiralité du foldamère.

Les informations tirées de cette expérience ne sont pas très résolues, elles ont néanmoins le mérite de préciser les fonctions participant à l'induction de chiralité du foldamère par la protéine. Pour obtenir des informations à haute résolution des interactions foldamère/protéine, nous avons tenté de cristalliser ce complexe (**8/HCA-II**) afin d'en déterminer la structure par diffraction des rayons X.

### 3.3. Structure du complexe **8/HCA-II**

#### 3.3.1. : *Cristallisation du complexe 8/HCA-II*

Des expériences de criblage de conditions de cristallogénèse (Crystal Screen, SaltRx, JCSG+) ont permis de dégager plusieurs conditions donnant de petits cristaux du complexe poussant dans un précipité (**Figure III.21.A**), dont la condition D9 du Crystal Screen (0,2 M ZnAc, 18 % PEG 8 000 dans 100 mM cacodylate de sodium, pH 6,5). Après optimisation par diffusion de vapeur en goutte suspendue, des cristaux de 100 x 50 x 50 µm ont pu être obtenus (**Figure III.21.**). Les conditions ayant permis l'obtention des cristaux les plus gros et de plus jolis faciès contiennent 0,2 mM ZnAc, 24 % PEG 8 000, 3 mM NaN<sub>3</sub> dans 100 mM cacodylate de sodium, pH 7,15. Dans les expériences précédentes (*screens* et optimisation), le rapport foldamère : protéine est maintenu constant. Sa valeur est de 2,4. La solution de protéine utilisée est à une concentration de 0,3 mM. Après congélation suivant la procédure décrite précédemment, nous avons testé les cristaux obtenus par diffraction des rayons X.

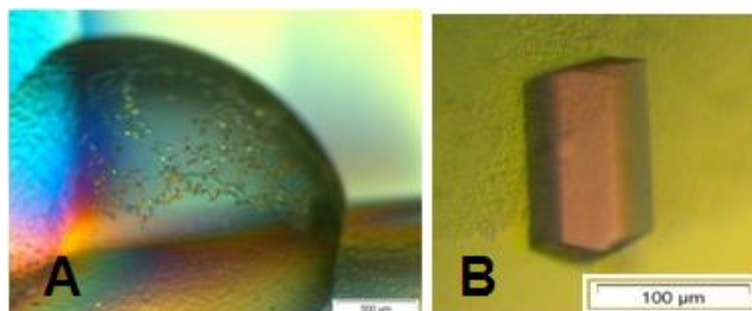


Figure III.21. : Cristaux du complexe 8/HCA-II.

A : Image d'une goutte de l'expérience de criblage (goutte D9 du kit Crystal Screen). B : Cristal obtenu et utilisé pour les expériences de diffraction des rayons X.

### 3.3.2. Structure du complexe 8/HCA-II

#### 3.3.2.1. Collecte et analyse des données

La collecte des intensités diffractées a été conduite sur la ligne de lumière ID23-2 à l'ESRF (Grenoble, France). Les 153 images ont été prises avec un temps d'exposition de 1 s et un incrément angulaire de 1,1 °. Les données ont ensuite été traitées à l'aide du programme XDS (Kabsch 2010). Le complexe cristallise dans le groupe d'espace  $P2_1$ , dans une maille cristalline ayant un volume deux fois plus important que celui fréquemment observé pour la HCA-II. Le coefficient de Matthews (Kantardjieff et Rupp 2010) est de 2,46 Å<sup>3</sup>/Da, et le pourcentage de solvant dans la maille est de 49,6 %. Ces données sont obtenues en considérant la présence de deux complexes foldamère/protéine indépendants dans la maille. La structure du complexe a été résolue par remplacement moléculaire avec comme modèle le fichier pdb de code 3KS3. Les programmes utilisés pour la résolution du problème des phases, ainsi que pour l'affinement sont listés dans la partie IV. § 4.2.2.3. Les statistiques de la collecte et de l'affinement sont présentées dans le **Tableau III.10**. La structure obtenue a une résolution de 2,15 Å (code PDB : 4LP6). Les coordonnées atomiques obtenues ont été validées à l'aide du programme Molprobit (Chen *et al.* 2010).

## 8/HCA

Statistiques de la collecte	
Groupe d'espace	$P2_1$
Paramètre de maille (Å, °)	$a = 45,960, b = 84,170, c = 76,700, \alpha = \gamma = 90, \beta = 97,180$
Résolution (Å)	41,5 - 2,15 (2,204 - 2,150)
$R_{\text{merge}}$	0,062 (0,837)
$I/\sigma(I)$	18,00 (2,21)
Complétude	99,7 (98,2)
Multiplicité	6,3 (6,2)
Statistiques en fin d'affinement	
Résolution (Å)	41,5 - 2,15 (2,204 - 2,148)
No. de réflexions	31628 (2166)
$R_{\text{work}}$	0,1693 (0,3040)
$R_{\text{free}}$	0,2329 (0,3890)
R.m.s.d	
longueurs de liaison (Å)	0,017
angles de liaison (°)	1,929
Hors Ramachandran	0 %
Ramachandran favorable	94,51 %
Facteur d'agitation thermique moyen (Å <sup>2</sup> )	
protéine	51,54
composé 8	46,30
H <sub>2</sub> O	58,28
Nombre de molécules	
H <sub>2</sub> O	175
glycérol	3
composé 8	2

Tableau III. 10. : Statistiques du traitement des données et de l'affinement.

Les données entre parenthèses correspondent à la couronne de plus haute résolution.

### 3.3.2.2. Description générale, dimérisation du complexe

Comme dans le cas de la structure modèle déterminée à résolution atomique (code PDB : 3KS3), les trois premiers résidus de chaque chaîne protéique, A et B, de la maille, apparaissent désordonnés. Tous les autres résidus ont pu être positionnés. Aussi, le modèle final montre-t-il une continuité de la carte de densité électronique (2mFo-DFc) contournée à 1,0  $\zeta$  tout le long des chaînes principales. Les molécules d'eau ont été positionnées d'abord automatiquement (Coot : Emsley *et al.* 2010) puis vérifiées manuellement grâce à la carte de densité électronique mFo-DFc contournée à 3,0  $\zeta$ , et grâce à la présence de liaisons hydrogène (distances comprises entre 2,4 et 3,6 Å). Dans la structure, trois molécules de glycérol, provenant de l'étape de cryo-protection, ont également été positionnées.

La structure fait apparaître un dimère régi par une symétrie non cristallographique de type C2. Cet agencement (**Figure III.22.**) induit de nouvelles interactions protéine/protéine, foldamère/protéine et foldamère/foldamère.

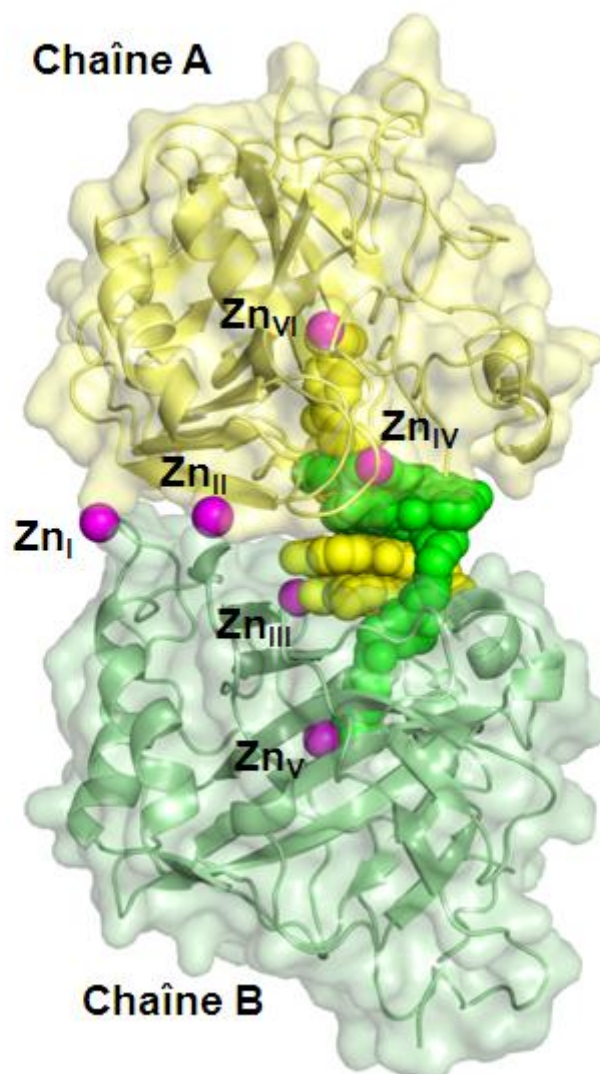
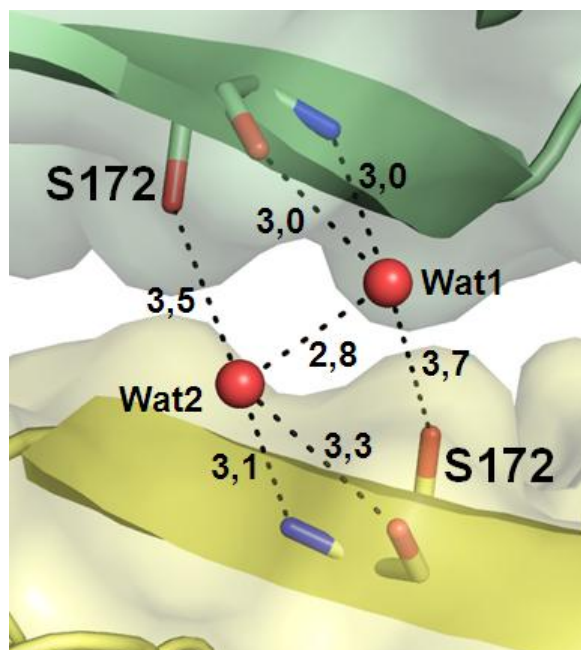


Figure III.22. : Structure cristalline du complexe 8/HCA-II.

Les deux chaînes protéiques sont différenciées par la couleur de leur surface : chaîne A en jaune clair, chaîne B en vert clair. Les composés 8 sont représentés par des sphères. Les atomes de zinc établissant des contacts protéine/protéine ( $Zn_I$  et  $Zn_{II}$ ) et 8/protéine (tous les autres) sont représentés par des sphères violettes.

### 3.3.2.3. Contacts protéine/protéine

Au sein d'un dimère, les contacts protéine/protéine impliquent deux molécules d'eau (Wat1 et Wat2) reliant les résidus S172 de chacune des deux protéines par des liaisons hydrogène (**Figure III.23.**), ainsi que deux sphères de coordination centrées sur des cations  $Zn^{2+}$  ( $Zn_I$  et  $Zn_{II}$  dans les **Figures III.22.** et **III.24.**).

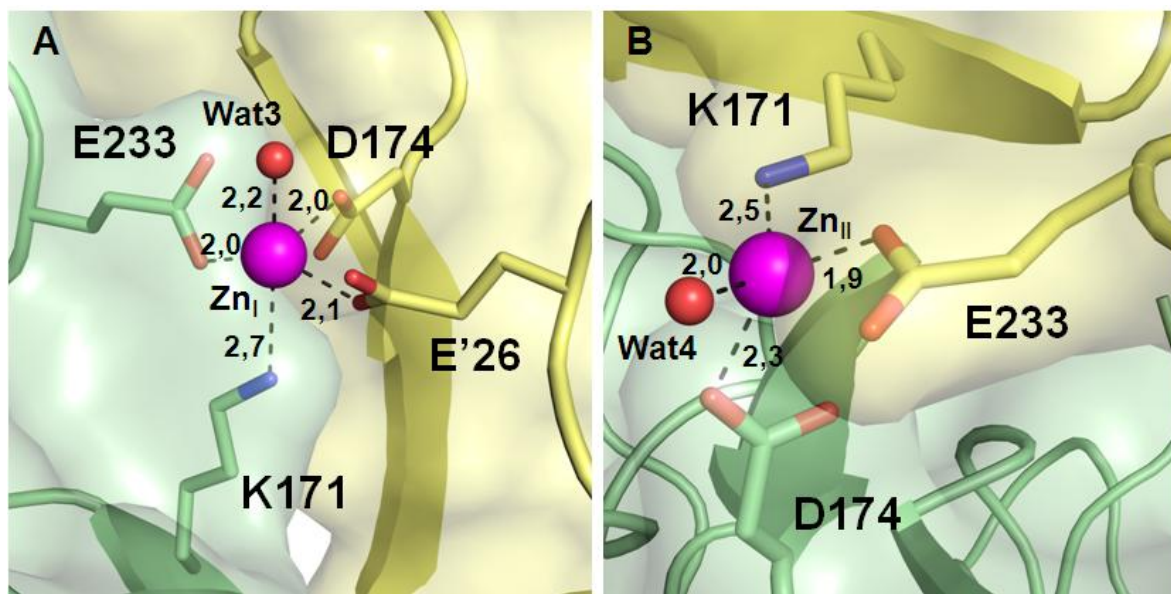


**Figure III.23. : Contacts protéine/protéine intra-dimère au niveau des résidus S172.**

Les sérines S172 sont entièrement représentées en bâtonnets pour mieux identifier les liaisons hydrogène en pointillés noirs. Les distances sont en angström.

Ce réseau de liaison hydrogène relie la chaîne principale et la chaîne latérale des sérines S172 avec deux molécules d'eau. La distance séparant Wat1 de l'oxygène de la chaîne latérale de S172 appartenant à la chaîne A (en jaune) est légèrement supérieure à la longueur habituelle d'une liaison hydrogène ( $< 3,6 \text{ \AA}$ ) mais l'agitation thermique de cet atome et de la molécule d'eau peut réduire cette distance finalement très proche d'une liaison hydrogène forte.

Les ions  $\text{Zn}^{2+}$ , provenant des conditions de cristallisation, sont essentiels à la formation des cristaux, puisqu'aucun cristal n'a pu être obtenu en absence de  $\text{Zn}^{2+}$ , ou en présence d'autres cations métalliques tels que  $\text{Mg}^{2+}$ . Ces cations sont impliqués dans les contacts protéine/protéine.  $\text{Zn}_\text{I}$  et  $\text{Zn}_\text{II}$  ont chacun une sphère de coordination incluant K171 et E233 d'une chaîne protéique donnée et D174 de l'autre (**Figure III.24.A et B**). Notons que  $\text{Zn}_\text{I}$  possède un ligand supplémentaire que  $\text{Zn}_\text{II}$  n'a pas : E26, appartenant à une chaîne protéique A symétrique voisine (opération de symétrie  $1+x, y, z$ ) (**Figure III.24.A**).



**Figure III.24. : Interactions protéine/protéine centrées autour des atomes de zinc Zn<sub>I</sub> (A) et Zn<sub>II</sub> (B).**  
Les distances mentionnées sont en angström.

La géométrie de coordination des atomes de zinc correspond parfaitement à celles rencontrées dans la littérature (Albert *et al.* 1998). Zn<sub>I</sub> et ses ligands adoptent une conformation pyramidale à base carrée. La seconde sphère de coordination (Zn<sub>II</sub>) adopte, quant à elle, une conformation tétragonale. Les résidus impliqués sont ceux généralement observés.

Au sein du cristal, les interactions inter-dimère sont relativement limitées. Elles font intervenir des réseaux peu développés de liaisons hydrogène et des sphères de coordination centrées autour d'atomes de zinc dont celui précédemment discuté (Zn<sub>II</sub>). Les valeurs élevées des facteurs d'agitation thermique observées pourraient être expliquées par ce faible nombre d'interactions (B moyen des chaînes protéiques : 51,5 Å).

#### 3.3.2.4. Les inhibiteurs interagissent avec leur site actif

Comme nous l'attendions, les parties inhibitrices de chacun des composés **8** sont liées à la protéine via la fonction sulfonamide coordonnée au zinc du site actif (Zn<sub>V</sub> et Zn<sub>VI</sub>, **Figure III.22.**).

En plus de cette interaction, les atomes du cycle benzénique proche du sulfonamide établissent des contacts hydrophobes avec L197. Le cycle benzyle interagit avec les résidus F130, V134 et L203, ainsi qu'avec le squelette quinoline de l'unité Q<sup>Asp</sup> (interaction de type  $\pi$ -stacking) et la chaîne latérale de l'unité Q<sup>Leu</sup> du second foldamère (**Figure III.25.**).



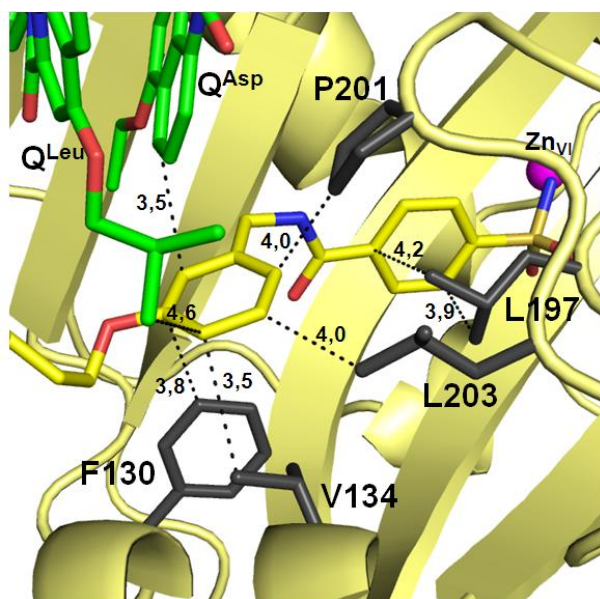


Figure III.25. : Interaction de la partie inhibitrice du composé 8 avec la protéine.

Le composé 8 en bâtonnets jaune est ancré dans le site actif de la protéine en ruban jaune, le second composé 8 est en bâtonnets vert. Les distances sont en angström.

La modification de l'inhibiteur initial par greffage de l'espaceur et ancrage du foldamère ne semble pas affecter la fixation de l'inhibiteur au site actif. La géométrie d'interaction entre les atomes du groupement sulfonamide et l'atome de zinc du site actif est inchangée. Les facteurs d'agitation thermique des atomes du composé **8** présents au site actif, et obtenus en considérant un facteur d'occupation égal à 1 (occupation systématique du site actif), sont comparables à ceux des atomes de la protéine au proche voisinage de ce site. Cette observation vient conforter les données déduites des expériences de SPR. Une superposition des composés **1** et **8** complexés avec la HCA-II est présentée **Figure III.26**. Seule l'orientation du benzyle terminal est légèrement modifiée puisque les deux atomes d'oxygène sont à 3,6 Å l'un de l'autre.

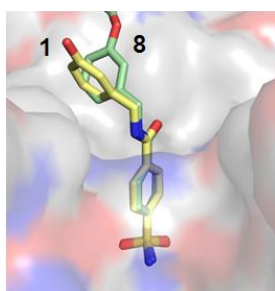


Figure III.26. : Superposition des molécules 1 et 8 au niveau du site actif de l'enzyme.



Les deux foldamères adoptent une conformation en hélice droite (P), en accord avec les résultats de dichroïsme circulaire induit. Les coefficients d'agitation thermique des deux chaînes protéiques et des deux foldamères sont sensiblement différents entre eux. La protéine la plus agitée (chaîne A en jaune :  $57,4 \text{ Å}^2$ ) est associée à un foldamère ordonné ( $45,3 \text{ Å}^2$ ). Ce dernier interagit avec la protéine la plus ordonnée (chaîne B en vert :  $45,4 \text{ Å}^2$ ), le second foldamère étant plus agité ( $47,3 \text{ Å}^2$ ). Le foldamère dont l'agitation thermique est la plus élevée ne nous révèle pas toutes ses chaînes latérales, seules les chaînes leucine et hydroxyle des deux premières unités sont ordonnées au sein de la structure. La position des sites actifs se faisant face permet des interactions foldamère/foldamère qui participent à la dimérisation.

### 3.3.2.5. Les foldamères interagissent entre eux ...

Des interactions de type  $\pi$ -stacking maintiennent les premières unités  $Q^{\text{Leu}}$  de chaque foldamère parallèles et à une distance inter-planaire moyenne de  $3,46 \text{ Å}$ . Ces dernières ne sont pas superposables puisqu'une rotation d'environ  $90^\circ$  existe entre elles (**Figure III.27.**). Les deux foldamères sont presque parfaitement alignés. Les positions atomiques des atomes constituant les noyaux quinolines de chaque foldamère ont permis de calculer, grâce au programme Mercury (Edington *et al.* 2008), l'équation de l'axe de chacun des foldamères. L'angle entre ces axes est voisin de  $7^\circ$ .

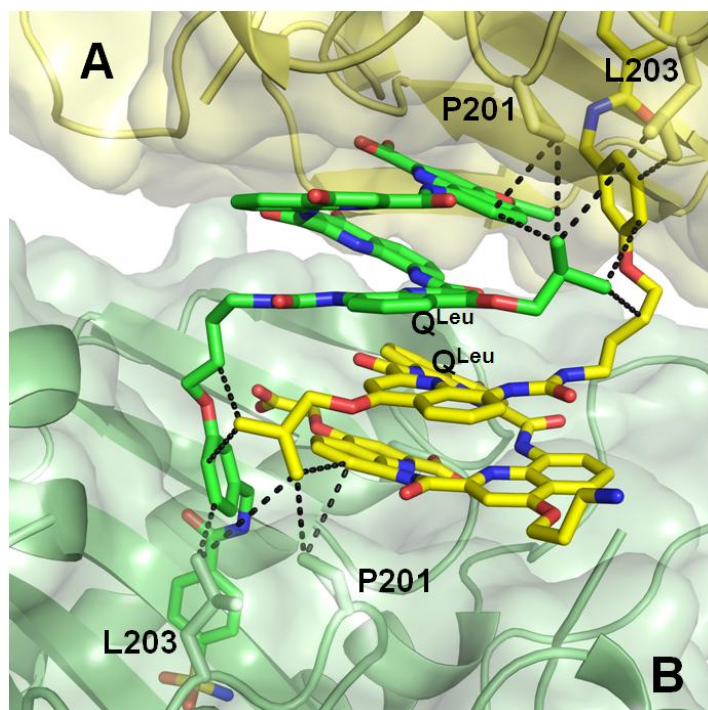
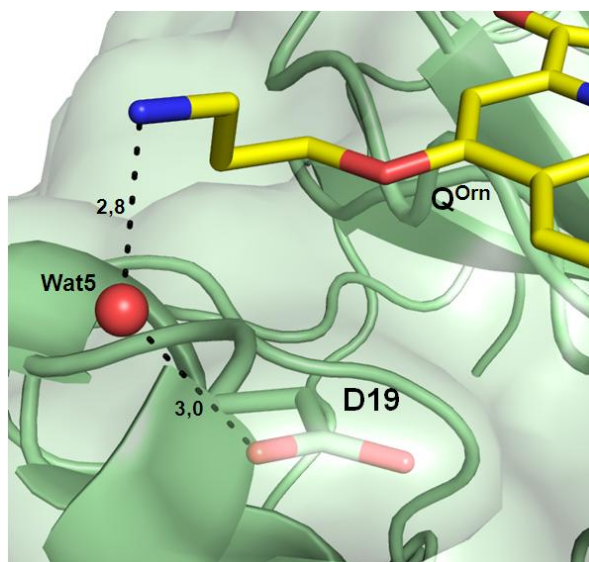


Figure III.27. : Interactions foldamère/foldamère et foldamère/protéine impliquant  $Q^{\text{Leu}}$ .

### 3.3.2.6. ... et avec les protéines

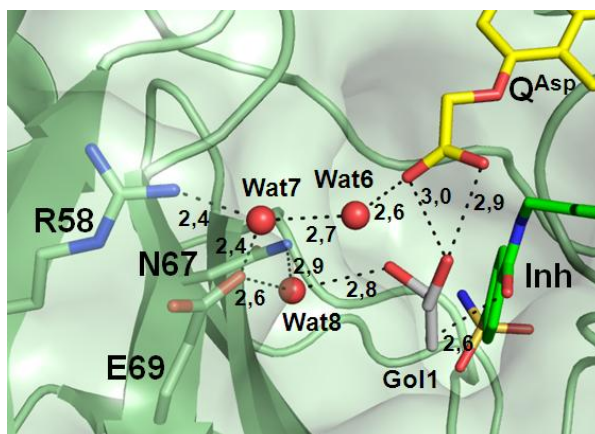
La première interaction foldamère/protéine le long du foldamère consiste en des interactions de type hydrophobe. La chaîne latérale de Q<sup>Leu</sup> se trouve à l'interface d'un environnement hydrophile (le solvant) et d'un domaine hydrophobe. Ce dernier est constitué entre autres de l'espaceur alkyle, du cycle benzyle du second composé **8**, du squelette quinoline de Q<sup>Asp</sup>, mais également des résidus P201 et L203 (**Figure III.27.**) Les interactions de type  $\pi$ -stacking décrites précédemment auxquelles s'ajoutent les deux domaines hydrophobes sur lesquels reposent les chaînes latérales leucine, confèrent aux interactions foldamère-foldamère un caractère hydrophobe très marqué et étendu (**Figure III.27.**).



**Figure III.28. : Interaction entre Q<sup>Orn</sup> et D19.**

Les distances indiquées sont en angström.

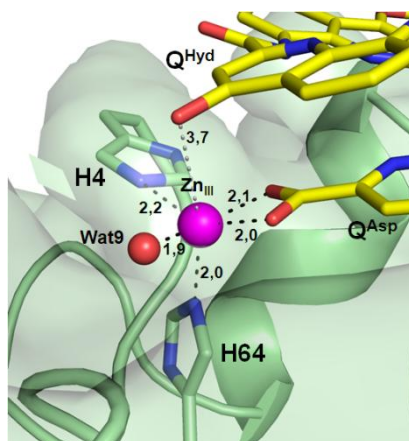
Lorsqu'elles sont observées, les chaînes latérales hydrophiles (ornithine et aspartate) du foldamère interagissent avec la protéine opposée par des liaisons hydrogène via des molécules de solvant (eau ou glycérol). Ainsi la chaîne latérale de Q<sup>Orn</sup> établit une interaction avec D19 de la protéine opposée via une molécule d'eau (**Figure III.28.**).



**Figure III.29. :** Réseau de liaisons hydrogène avec la chaîne latérale de Q<sup>Asp</sup>.

Ce réseau inclue des molécules d'eau (Wat) et une molécule de glycérol (Gol). Les distances rapportées sont en angström.

La chaîne latérale de Q<sup>Asp</sup> interagit quant à elle avec les résidus N67, E69, R58 de la protéine opposée, ainsi qu'avec un atome d'oxygène du second inhibiteur par l'intermédiaire de liaisons hydrogène via des molécules d'eau et une molécule de glycérol notée Gol1 (**Figure III.29.**).



**Figure III.30. :** Interaction 8/HCA-II médiée par un ion Zn<sup>+2</sup>.

Les distances rapportées sont en angström.

La partie C-terminale de chacun des deux foldamères interagit avec la HCA-II opposée, grâce à des liaisons de coordination avec un Zn<sup>2+</sup> (Zn<sub>III</sub> et Zn<sub>IV</sub> sur la **Figure III.22.**) dont la sphère de coordination est pyramidale à base carrée. Les ligands de ces atomes de zinc sont la fonction carboxylate terminale de la dernière unité quinoline du foldamère, l'atome Ne des histidines 4 et 64 et une molécule d'eau (**Figure.III.30.**). L'oxygène de la chaîne latérale hydroxyle de Q<sup>Hid</sup> est diamétralement opposé à l'atome Ne du résidu H64, toutefois la

distance qui le sépare de l'atome de zinc exclue de le considérer comme un ligand du cation (3,7 Å au lieu d'environ 2,1 Å).

Les chaînes protéiques A et B des 2 molécules indépendantes adoptent le même repliement. Le r.m.s.d moyen obtenu en superposant les 257 C $\alpha$  de chacune des deux chaînes protéiques est de 0,23 Å. L'écart maximum (0,93 Å) est observé au niveau des C $\alpha$  du résidu K76. Ce repliement est semblable à celui observé pour la protéine native : les r.m.s.d moyens calculés (sur les positions des 257 C $\alpha$  précédents) pour chaque chaîne (A et B) comparativement au modèle (code PDB : 3KS3) sont respectivement de 0,33 Å et 0,31 Å. L'écart maximal est de 1,14 Å pour la chaîne A. Il concerne l'écart de positions entre C $\alpha$  de A77. Celui de la chaîne B est de 0,95 Å (P236). Notons que les résidus K76, A77 et P236, pour lesquels les maximum de r.m.s.d. sont observés, sont tous situés en surface de la protéine et à grande distance du site actif. Ainsi, les nombreuses interactions entre l'inhibiteur, l'espaceur, le foldamère, et la protéine ne modifient-elles pas de façon significative la conformation de la chaîne principale protéique.

## Conclusion

L'anhydrase carbonique se révèle être un bon modèle pour valider la stratégie que nous utilisons pour reconnaître des surfaces protéiques par des foldamères.

Les expériences de SPR ont démontré qu'un foldamère ancré à un inhibiteur de haute affinité pour la protéine ne modifiait pas sensiblement cette affinité. Ces résultats ouvrent la voie à de premières études de reconnaissance de surface. Il est désormais possible de maintenir, à proximité de la surface de la HCA-II, un foldamère tout en lui laissant la possibilité ou non, via la flexibilité d'un espaceur, d'interagir avec elle.

Les expériences menées par CD induit confirment, quelque soit la méthode de synthèse du foldamère (en solution ou sur support solide) que ce dernier subit une induction d'hélicité en présence de la protéine. Le foldamère est sensible à la proximité de cette dernière et semble donc capable d'interagir avec elle. Peut-être plus important encore, l'intensité du signal induit dépend du foldamère utilisé, ce qui nous rassure sur le fait que la fonctionnalisation de celui-ci est un paramètre important, susceptible de générer ou non les interactions attendues.

Ces résultats nous ont conduits à envisager l'étude structurale du complexe 8/HCA-II (complexe présentant le signal de CD induit le plus important) afin de vérifier l'existence

d'interactions spécifiques et de les caractériser. Des cristaux ont été obtenus et la structure du complexe a été décrite à la résolution de 2,15 Å par diffraction des rayons X. Il s'agit, à notre connaissance, de la première structure d'un foldamère de quinolines complexé à une protéine. Cette structure ne montre pas de façon évidente les interactions initialement espérées c'est-à-dire des interactions franches entre un foldamère donné et la protéine à laquelle il est ancré. Les interactions trouvées sont multiples (interactions de type protéine/protéine, foldamère/protéine, et foldamère/foldamère) principalement régies par des atomes de zinc pontant chaque partenaire l'un à l'autre.

Cette structure est toutefois importante. Elle valide tout d'abord la stratégie d'ancrage utilisée. Son analyse attentive apporte différents indices permettant, nous l'espérons, de renforcer des interactions intra-monomères au profit de la dimérisation observée. Pour y parvenir, nous devons préalablement définir si cette dimérisation existe en solution et si elle est induite par la présence d'atomes de zinc issus des conditions de cristallisation. Rappelons en effet qu'aucun cristal n'a pu être obtenu pour le moment en absence de zinc ou en présence de cations divalents autres, tels que  $Mg^{2+}$ . Des expériences de diffusion de lumière, de RMN, voire plus simplement de chromatographie d'exclusion de taille, devraient nous permettre de répondre à ces premières questions.

Les résultats obtenus en CD induit avec les foldamères permettant de valider l'importance de chaque chaîne latérale s'expliquent difficilement à la vue de la structure 3D. Quelques questions restent en suspens. Pourquoi, par exemple, la modification de la chaîne ornithine en simple atome d'hydrogène n'a-t-elle aucun effet sur le spectre de dichroïsme circulaire alors que toutes les autres modifications en ont un ? Ces observations trouvent difficilement une explication au regard de la structure 3D. Rappelons que les interactions trouvées entre chaînes latérales d'un foldamère et atomes de la protéine ne semblent pas prépondérantes (liaisons hydrogène via des molécules de solvant). L'absence de corrélations claires entre données d'ICD et données structurales ne peut-elle pas indiquer que la dimérisation n'a pas lieu en solution mais résulte du passage à l'état cristallin ?

Eviter la dimérisation nécessite de savoir quels sont les paramètres qui la favorisent. Quelle est l'importance des atomes de zinc ? De simples mutations des acides aminés impliqués dans les contacts protéine-protéine permettraient-elles de s'affranchir de ces interactions ? Une protection de la fonction carboxylate en C-ter du foldamère et/ou l'allongement de ce dernier par ajout de sous-unités quinolines déstabiliseraient-ils les interactions entre un foldamère donné et la protéine appartenant à l'autre monomère ?

# Conclusion générale et perspectives

Les interactions protéine - protéine sont au cœur de tous les grands processus biologiques. Elles sont de ce fait des cibles thérapeutiques très largement étudiées. La caractérisation des mécanismes biologiques dans lesquels elles interviennent, la mise en évidence de nouvelles interactions et l'étude de leur possible inhibition mobilisent de très nombreuses équipes scientifiques. Diverses stratégies (utilisation d'anticorps, de protéines miniatures, de plus petites molécules, ...) ont été développées afin d'inhiber ces interactions et ainsi de stopper le déroulement d'un processus biologique impliqué dans une pathologie donnée. Toutes ces approches partagent un point commun : concevoir une molécule antagoniste (squelette protéique, nucléotidique ou de synthèse) plus affine pour une surface protéique donnée que la protéine partenaire impliquée dans l'interaction que l'on souhaite inhiber.

De nouvelles classes de molécules, les foldamères, semblent particulièrement pertinentes pour inhiber les interactions protéine – protéine. Nous avons rappelé dans l'introduction de ce mémoire les principales propriétés de ces molécules de synthèse. Leur structure 3D est parfaitement prévisible. Le repliement adopté est celui rencontré dans les structures secondaires de protéines. Ces molécules sont en outre aisément fonctionnalisables (car élaborées à partir d'une bibliothèque de monomères portant une ou plusieurs chaînes latérales protéinogéniques) et insensibles à la dégradation protéolytique.

Les travaux réalisés dans ce mémoire ont porté sur des foldamères d'oligoamides de quinoline, molécules adoptant une structure tridimensionnelle hélicoïdale. Le but de l'étude consistait à montrer que ces composés étaient susceptibles de reconnaître (et donc d'interagir avec) une surface protéique prédéfinie. Ce travail s'est inscrit, nous l'avons préalablement mentionné, dans un projet européen de type IAPP « Industry – Academia Partnerships and Pathways » intitulé FOLDAPPI (FOLDamers Against Protein – Protein Interactions) incluant principalement chimistes, biochimistes et structuralistes issus du monde académique et privé. Il a donc, bien évidemment, bénéficié de l'expertise et de la multidisciplinarité des différents groupes participant à ce programme. La stratégie (« *Structure Based Design* ») adoptée pour la conception des foldamères utilisés consistait à modéliser et concevoir *in silico* la séquence d'un composé susceptible d'interagir avec la surface protéique cible, puis à optimiser cette séquence en fonction des données

acquises lors d'études biophysiques (CD, SPR, diffraction des rayons X) cherchant à caractériser les interactions espérées. Ma contribution à ce travail consistait à mener les études biophysiques. Tous les foldamères utilisés ont été préalablement synthétisés dans l'équipe de chimistes dirigée par Ivan Huc. Le choix des protéines cible était en outre déjà décidé lorsque j'ai débuté ces études.

Deux protéines ont été retenues pour mener à bien cette étude : l'interleukine 4, protéine impliquée dans de nombreux processus immunitaires (allergie, asthme, ...) et une enzyme, l'anhydrase carbonique humaine de type II. Le choix de ces deux protéines (dont nous avons rappelé les principales données issues de la bibliographie) était dicté par la parfaite connaissance de leurs propriétés structurales et fonctionnelles, maintes fois décrites et particulièrement adaptées aux études envisagées.

Il nous est apparu important, lors de ces premières études, d'optimiser les chances d'interaction entre le foldamère et la protéine. Il fallait pour cela maintenir le foldamère à proximité de la surface à reconnaître tout en lui laissant un degré de liberté suffisamment grand afin de ne pas imposer ces interactions. L'originalité du travail présenté se situe sans doute à ce stade. Nous avons choisi d'ancrer un foldamère à la protéine (ou au site actif de l'enzyme) via un espaceur. Dans le cas de l'interleukine 4, c'est la mutation d'une sérine en cystéine qui permet l'ancrage d'un foldamère modifié, grâce à une liaison thioéther (utilisation d'un groupement maléimide) ou à un pont disulfure (utilisation d'une fonction sulfhydryle activée). En ce qui concerne l'anhydrase carbonique, la stratégie d'ancrage est basée sur l'utilisation d'un inhibiteur modifié pour lequel la fixation de l'espaceur et du foldamère n'altère que faiblement la constante d'affinité qui reste de l'ordre du nanomolaire.

Pour les deux protéines étudiées nous avons pu vérifier que l'hypothèse de départ, à savoir qu'un foldamère peut interagir avec une surface protéique, était plausible. Quelque soit la protéine considérée, IL-4 ou HCA-II, nous avons montré par dichroïsme circulaire induit que le foldamère ancré à la protéine subissait un changement d'hélicité et qu'il adoptait ainsi une chiralité préférentielle. Des interactions existeraient donc entre partenaires. Les caractériser nécessite une description structurale des complexes obtenus. Nous n'y sommes pas encore parvenu dans le cas de l'interleukine 4 compte tenu, nous l'avons vu, de difficultés rencontrées concernant le repliement de la protéine extraite des corps d'inclusion. Toutefois, dans le cas de l'anhydrase carbonique humaine de type II, une première structure a été déterminée à 2,15 Å de résolution par diffraction des rayons X. Cette structure entre protéine et foldamère d'oligoamides de quinolines est, à notre connaissance, la première décrite. Les coordonnées atomiques ont été déposées à la Protein Data Bank sous le

code : 4LP6. L'ensemble du travail effectué sur cette enzyme a donné lieu à publication. L'article intitulé "Structure at 2.1 Å resolution of a helical aromatic oligoamide foldamer-protein complex" vient d'être accepté par la revue *Angewandte Chemie*. Une copie des épreuves est jointe.

Les analyses de spectrométrie de masse après digestion par la trypsine de l'interleukine 4 mutée ont montré que seule une fraction de la protéine était correctement repliée. Pourquoi ? Est-ce dû à un temps de stockage de quelques mois pendant lequel la protéine était maintenue à -20°C ? Cela reste possible compte tenu du caractère instable de la protéine. Rappelons que les étapes d'extraction et de purification ont dû être conduites à 4°C. Bien que l'étape de déglutathionylation ait été réalisée à température ambiante, cette étape ne pourrait être mise en cause au vu de la structure obtenue et des résultats de dichroïsme circulaire. Il semble nécessaire de préciser l'étape à laquelle a lieu la dénaturation partielle. Dans le cas où des résultats similaires sont obtenus sur un lot de protéine fraîchement produite, il sera nécessaire de modifier la méthode de production de façon à obtenir un rendement suffisant en protéine soluble. Quoiqu'il en soit, il est nécessaire de poursuivre le travail déjà entrepris et de résoudre la structure de l'adduit ayant montré un fort signal positif de dichroïsme circulaire induit.

L'étude menée sur HCA-II a débuté avec des foldamères synthétisés en solution. Les premiers résultats de dichroïsme circulaire montraient une possible interaction. Il nous a fallu toutefois reprendre ces études sur de nouveaux composés, suite à l'application d'une nouvelle méthode de synthèse plus performante : la synthèse sur support solide a été immédiatement adoptée par les chimistes.

La structure 3D obtenue montre une dimérisation inattendue du complexe. Quand apparaît cette dimérisation ? Quelles sont les interactions qui la gouvernent ? Peut-on affaiblir ou au contraire renforcer certaines d'entre elles ?

Des expériences de diffusion dynamique et statique de la lumière (DLS et SLS), de chromatographie d'exclusion de taille et de RMN devraient nous indiquer si des dimères existent en solution. L'analyse de la structure montre l'importance des cations métalliques dans les interactions protéine - protéine et foldamère - protéine. Quel est le rôle du zinc dans la dimérisation ? Des expériences de DLS conduites en présence et en absence de zinc dans la solution permettraient de répondre à cette question.

Les interactions foldamère - protéine observées dans la structure sont inattendues. Elles mettent en contact le foldamère d'un monomère donné avec la protéine du second monomère.



Rappelons que nous recherchions des interactions entre un foldamère et la protéine auquel il est ancré. Quelles modifications doit-on effectuer pour les observer ? La simple mutation de résidus impliqués dans les contacts protéine - protéine permettrait-elle de s'affranchir de la dimérisation et d'observer les interactions espérées ? Doit-on rallonger le foldamère et/ou protéger la fonction carboxylate en position C-terminale afin d'affaiblir les interactions entre un foldamère et la protéine opposée ? Notons que l'ajout de sous-unités quinoline judicieusement fonctionnalisées augmenterait le nombre potentiel de contacts au sein d'un monomère.

La structure 3D obtenue nous montre, s'il en était besoin, que nous sommes encore loin de la reconnaissance spécifique d'une surface protéique prédéfinie. Toutefois, des progrès notables ont été réalisés sans que l'on ait, pour le moment, réellement optimisé les séquences des foldamères étudiés. L'ensemble de ce travail montre qu'un foldamère de quinoline est capable d'interagir avec une surface protéique de façon telle que cette dernière induise une hélicité particulière (hélice droite). L'expertise acquise lors de ce travail doit pouvoir être utilisée sur de nouvelles cibles à visée thérapeutique présentant, comme l'interleukine 4, des surfaces d'interaction parfaitement définies.

---

## Chapitre IV

### Matériels et méthodes



# 1. Biologie moléculaire et cellulaire

## 1.1. Souches bactériennes utilisées

La production de l'IL-4 mutée s'effectue dans la souche *Escherichia coli* Rosetta (DE3) [ $F^-$ , *ompT*, *hsdS<sub>B</sub>*( $r_B^-m_B^-$ ) *gal dcm* (DE3) pLysSRARE] résistante au chloramphénicol.

La HCA-II est exprimée dans la souche *Escherichia coli* BL21 Gold (DE3) pLysS [ $F^-$ , *ompT* *hsdS*( $r_B^-m_B^-$ ) *dcm*<sup>+</sup> Tet<sup>r</sup> *gal*  $\lambda$ (DE3) *endA* Hte (pLysS Cam<sup>r</sup>)] résistante au chloramphénicol.

Les bactéries sont stockées sous forme d'aliquot à -80°C dans le milieu LB complémenté avec 50 % de glycérol.

## 1.2. Plasmides et séquences utilisés

Le plasmide pQKA contenant l'ADNc codant pour l'IL-4 (UniProtKB/Swiss-Prot P05112) portant la mutation S16C est utilisé pour l'expression bactérienne de la protéine. Il nous a été fourni par l'équipe du Dr. Thomas Mueller à Würzburg. La sélection est réalisée grâce à un gène de résistance à la kanamycine intégré au plasmide. L'expression est régulée par un promoteur inductible par le galactose et son analogue non hydrolysable l'IPTG.

Pour l'expression bactérienne de la HCA-II, le plasmide pETD11d contenant l'ADNc codant pour la protéine (UniProtKB/Swiss-Prot P00918) est utilisé. Ce plasmide a été construit par Sabrina Rousseau (Plateau technique de l'IECB) à Bordeaux. La sélection des bactéries est réalisée grâce à un gène de résistance à l'ampicilline présent dans le plasmide. L'expression est induite par le galactose ou son analogue non hydrolysable l'IPTG.

## 1.3. Quantification des fragments d'ADN

Les acides nucléiques peuvent être quantifiés en mesurant l'absorbance d'un échantillon à 260 nm. Le coefficient d'extinction molaire est de 0,020  $\mu\text{g/mL.cm}$  pour l'ADN double brin, soit une concentration de 50  $\mu\text{g/mL}$  pour une unité d'absorbance. Le rapport des absorbances à 260 nm et 280 nm permet de déterminer la pureté de l'échantillon. On considère que l'échantillon est uniquement composé d'acides nucléiques si le rapport

A260 / A280 est compris entre 1,8 et 2. S'il est inférieur, l'échantillon est contaminé par des protéines et s'il est supérieur à 2, l'échantillon présente une quantité importante d'ARN.

## 1.4. Mutagenèse dirigée

Les mutants ponctuels de la HCA-II ont été réalisés par Sabrina Rousseau (Plateau technique de l'IECB) à Bordeaux. Le mutant de l'IL-4 a lui été réalisé par l'équipe du Dr. Thomas Mueller à Würzburg.

## 1.5. Transformation bactérienne

Les bactéries sont décongelées dans de la glace, puis sont étalées sur une boîte contenant du milieu sélectif. Les cellules sont incubées 24 h à 37 °C, puis une colonie est reprise dans 50 µL d'eau milliQ. Après une incubation de 30 min à 0 °C en présence de 50 ng de plasmide, les cellules subissent un choc thermique à 42 °C pendant 45 s, pour être ensuite replacées à 0 °C pendant 2 min. Les cellules sont reprises avec du milieu SOC et placées à 37 °C sous agitation (180 rpm) pendant 1 h. Elles sont récoltées par centrifugation et finalement étalées sur une boîte contenant du milieu sélectif, et placées à 37 °C.

## 1.6. Milieux de culture

Les bactéries sont cultivées ou régénérées dans du milieu LB ou du milieu SOC sous agitation (180 rpm) à 37 °C. La composition de ces milieux est la suivante :

LB	pH 7,0	SOC	pH 7,0
NaCl	10 g/L	NaCl	0,5 g/L
Peptone	10 g/L	KCl	186 mg/L
Extrait de levure	5 g/L	MgSO <sub>4</sub> anhydre	2,4 g/L
+/- Agar	15 g/L	Peptone	20 g/L
+/- Ampicilline	100 µg/mL	Extrait de levure	5 g/L
+/- Chloramphénicol	200 µg/mL	Glucose	3,6 g/L
+/- Kanamycine	50 µg/mL		

Pour suivre la croissance cellulaire, des mesures d'absorbance à 600 nm sont effectuées.

## 2. Biochimie

### 2.1. Dosage des protéines

Les protéines sont quantifiées par spectrométrie d'absorption, en utilisant la loi de Beer-Lambert :

$$A = \epsilon.l.C$$

Avec  $A$  : l'absorbance ou DO sans unité

$\epsilon$  : le coefficient d'absorption molaire exprimé en  $M^{-1}.cm^{-1}$

$l$  : la longueur du chemin optique exprimée en cm

$C$  : la concentration exprimée en M

Les valeurs des coefficients d'absorption molaire ainsi que celles des points isoélectriques ont été calculées théoriquement à partir de la séquence primaire des protéines à l'aide du programme « ProtParam » disponible sur le serveur d'outils bioinformatiques ExPASy (<http://web.expasy.org/protparam>)

### 2.2. Electrophorèse des protéines par SDS-PAGE

L'électrophorèse des protéines sur gel de polyacrylamide en présence de SDS permet leur séparation en fonction de leur masse grâce à l'application d'un courant électrique (Læmmli 1970). Le SDS est un détergent ou surfactant anionique fort qui dénature les protéines en leur conférant des charges négatives. La quantité de SDS en interaction avec la chaîne protéique est directement proportionnelle à sa masse. Ainsi la charge de la protéine (fonction de sa masse) devient le critère principal de la séparation dans le gel.

Avant d'être déposé sur le gel, les échantillons sont solubilisés dans du tampon de charge Læmmli. Cette solution est dénaturante de par la présence de SDS et de  $\beta$ -mercaptoéthanol, composé permettant la réduction des ponts disulfures intra et inter-chaîne. Le glycérol est utilisé pour augmenter la densité de l'échantillon et ainsi faciliter son dépôt sur le gel.

Les protéines sont déposées sur un gel de concentration (4 % w/v acrylamide, 0,1 % w/v bis-acrylamide, 125 mM Tris-HCl pH 6,8, 0,15 % w/v SDS, 0,1 % w/v APS, 0,2 % v/v TEMED) dont le rôle est d'augmenter la résolution. Les protéines passent ensuite dans le gel

de séparation (10 à 15 % w/v acrylamide, 0,25 à 0,37 %w/v bis-acrylamide, 375 mM Tris-HCl pH 8,8, 0,15 % w/v SDS, 0,1 % w/v APS, 0,2 % v/v TEMED). La migration est effectuée dans un tampon de migration TGS (25 mM Tris-HCl, 200 mM Glycine, 0,15 % w/v SDS, pH 8,3) à une tension constante de 150 V.

## 2.3. Production et purification du mutant S16C de l'IL-4

Les bactéries sont mises en culture à 37 °C et sous agitation (180 rpm) dans un milieu LB contenant de la kanamycine jusqu'à atteindre une  $DO_{600nm}$  de 0,6. L'expression de l'IL-4 mutée est induite par l'ajout de 1 mM d'IPTG. Après 3 h à 37 °C, les cellules sont récoltées puis resuspendues dans 20 volumes de tampon STE (0,325 M saccharose, 10 mM Tris-HCl pH 8,0, 1 mM EDTA) complété avec 1 mM de DTT. Les cellules sont ensuite lysées grâce à un disrupteur (30 kpsi) (TS, Constant Systems LTD) ou par sonication (10 min entrecoupées toutes les 30 s d'une pause de 30 s, à 4 °C). La fraction insoluble (les corps d'inclusion notamment) est séparée de la fraction soluble par centrifugation (22 000 g, 30 min, à 4 °C), et lavée avec du tampon STE complété avec 1 mM de DTT d'abord puis avec du tampon STE complété avec 1 mM de DTT et 1 % de Triton X-100. Les corps d'inclusion sont conservés à -20 °C.

La protéine est extraite des corps d'inclusion grâce à une solution dénaturante (10 volumes) contenant 6 M GuCl, 0,1 M Tris-HCl pH 8,0, et 1 mM DTT. Pour renaturer la protéine, l'extraction protéique est diluée 20 fois dans du tampon de renaturation contenant 100 mM arginine, 100 mM Tris-HCl, pH 8,0, 5 mM EDTA, 5 mM de glutathion sous forme oxydée (GSSG), 2 mM de glutathion sous forme réduite (GSH), et incubée sous agitation à 4 °C, pendant 3 jours.

La solution est dialysée deux fois avec 10 volumes de PBS (137 mM NaCl, 2,7 mM KCl, 10 mM  $Na_2HPO_4$  et 2 mM  $KH_2PO_4$ , pH 7,4) à 4 °C pendant 3 à 6 h.

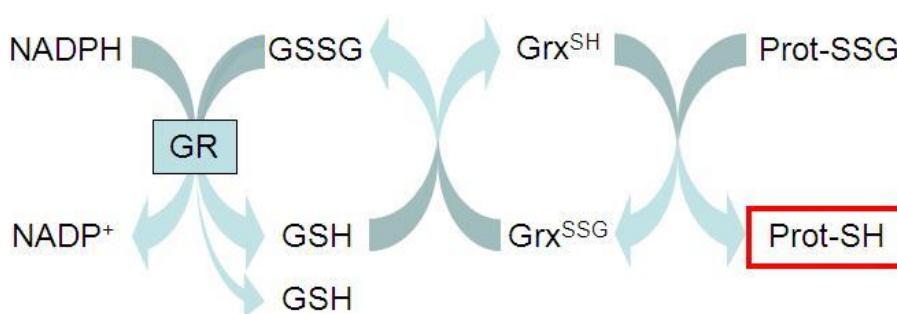
Le pH de l'extrait protéique est ajusté à 5,0 en ajoutant 0,01 volumes d'acétate d'ammonium 4 M, pH 5,0. Le mélange est clarifié par centrifugation (12 000 g, 30 min, à 4 °C) et filtré. Le filtrat est chargé sur une colonne MonoS HR 16/10 (GE Healthcare) préalablement équilibrée avec un tampon d'acétate d'ammonium 25 mM, pH 5,0. La protéine mutée est ensuite éluée par un gradient linéaire de NaCl (0 à 2 M). La fraction contenant l'IL-4 mutée est chargée sur une colonne de filtration sur gel (HiLoad Superdex 75 16/600)

(GE Healthcare) préalablement équilibrée avec un tampon contenant 20 mM  $\text{NaH}_2\text{PO}_4$ , pH 6,0 et 300 mM NaCl. Les fractions contenant la protéine sont réunies, concentrées jusqu'à 10 mg/ml, aliquotées et conservées à  $-20^\circ\text{C}$ .

A la fin des étapes de production et de purification, la protéine mutée porte une molécule de glutathion sur la cystéine libre introduite par la mutation. En effet pendant la renaturation de la protéine, un mélange de glutathion sous forme réduite et oxydée est utilisé, et la cystéine libre forme un pont disulfure avec une molécule de glutathion réduite (d'après Duppatla 2012).

## 2.4. Réduction enzymatique de l'adduit IL-4/Glutathion

La réaction s'effectue dans une solution contenant 100 mM  $\text{KH}_2\text{PO}_4$  pH 7,0, 2 mM EDTA, 6  $\mu\text{g}/\text{mL}$  de glutathion réductase, 0,2 mM NADPH, 0,5 mM de glutathion (forme réduite) et 1 à 2 mg/mL d'IL-4 mutée purifiée. La réaction est déclenchée par l'ajout de 6 unités de glutaredoxine (**Figure V.1.**). L'évolution de la réaction est suivie grâce à l'oxydation du NADPH, qui peut être mesurée par la décroissance de l'absorption à 340 nm ( $\epsilon_{340\text{nm}} = 6\,220\text{ M}^{-1}\cdot\text{cm}^{-1}$ ). Après 30 min, le mélange réactionnel est chargé sur une colonne de filtration sur gel HiLoad Superdex 75 16/600 (GE Healthcare) préalablement équilibrée avec un tampon contenant 20 mM  $\text{NaH}_2\text{PO}_4$  pH 6,0 et 300 mM NaCl. Les fractions contenant l'IL-4 mutée sont rassemblées, concentrées puis stockées à  $-20^\circ\text{C}$  (Duppatla 2012).



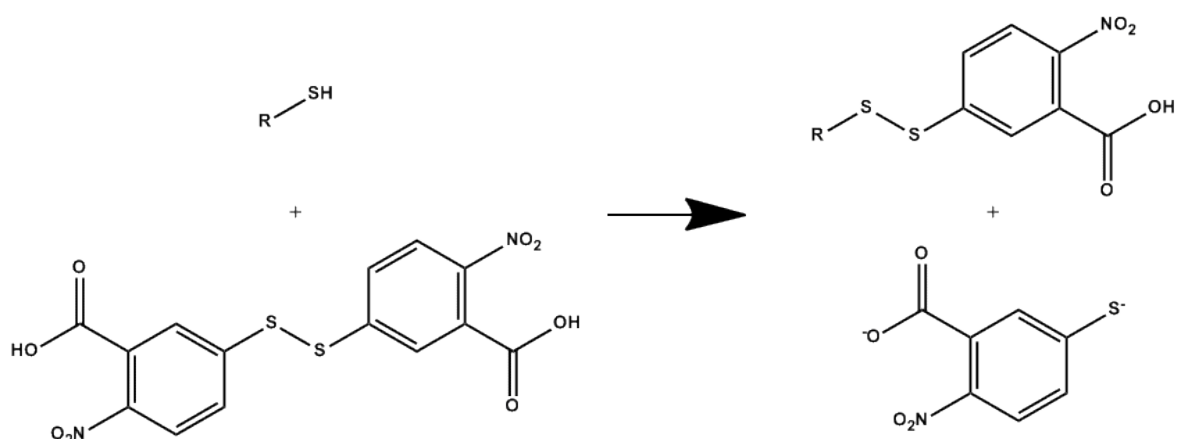
**Figure V.1. : Système enzymatique de réduction de l'adduit IL-4/Glutathion.**

Le glutathion porté par la protéine va être transféré à la Glutaredoxine qui va permettre de synthétiser le glutathion oxydé. Ce dernier sera finalement réduit par la Glutathion réductase avec la participation du NADPH. GR : Glutathion Réductase, GSSG/GSH Glutathion oxydé/réduit,  $\text{Grx}^{\text{SSG}}/\text{Grx}^{\text{SH}}$  : Glutaredoxine portant/ou non un glutathion, Prot-SSG/Prot-SH : Protéine portant/ou non un glutathion.



## 2.5. Quantification des cystéines libres par le réactif d'Ellman

Les thiols, et dans notre cas les cystéines libres des protéines sont quantifiés avec le réactif d'Ellman ou DTNB, en clivant le pont disulfure de la molécule pour donner le TNB qui, à pH neutre et alcalin, s'ionise en  $\text{TNB}^{2-}$  de couleur jaune (Ellman 1958 et Riddles *et al.* 1983). Cette réaction (**Figure V.2.**) est rapide et stœchiométrique.



**Figure V.2. :** Schéma de la réaction de dosage.

La fonction thiol libre va réagir avec le pont disulfure intramoléculaire du DTNB pour former un pont disulfure avec l'une des deux moitiés. Cette réaction est possible car la seconde moitié est un bon groupe partant stabilisé par résonance.

La quantification s'effectue dans un tampon 100 mM  $\text{KH}_2\text{PO}_4$  pH 7,4 et la mesure s'effectue à 412 nm. Le blanc est uniquement composé de tampon. Le contrôle permettant de calculer le coefficient d'extinction molaire du DTNB à 412 nm est lui composé de tampon et de DTNB. L'échantillon contient du DTNB, la protéine et du tampon. Le nombre de cystéines libres est calculé avec la formule suivante :

$$\text{Nombre de cystéines libres} = \frac{\left( \frac{A_{\text{Echantillon}}}{l} - C_{\text{DTNB}} \times \epsilon_{\text{DTNB}} \right) \times \frac{l}{(\epsilon_{\text{TNB}^{2-}} - \epsilon_{\text{DTNB}})} \times Fd_{\text{protéine}}}{C_{i \text{ protéine}}}$$

Avec  $l$  : la longueur du chemin optique exprimée en cm.

$C_{\text{DTNB}}$  : la concentration de DTNB dans la cuve exprimée en M.

$\epsilon_{DTNB}$  : le coefficient d'absorption molaire du DTNB à 412 nm exprimé en  $M^{-1}.cm^{-1}$  (calculé avec l'expérience contrôle  $\approx 250 M^{-1}.cm^{-1}$ ).

$\epsilon_{TNB^{2-}}$  : le coefficient d'absorption molaire du  $TNB^{2-}$  à 412 nm exprimé en  $M^{-1}.cm^{-1}$  (théoriquement égal à  $14\,150 M^{-1}.cm^{-1}$ ).

$Fd_{protéine}$  : le facteur de dilution de la protéine.

$C_{i\,protéine}$  : la concentration initiale de protéine exprimée en M.

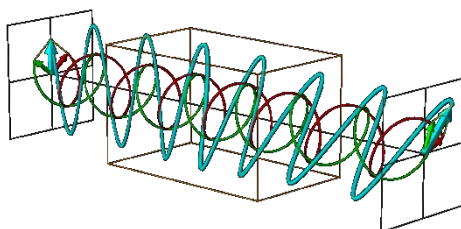
## 2.6. Production et purification de la HCA-II

La production et la purification de la HCA-II ont été réalisées par Sabrina Rousseau (Plateau technique de l'IECB) en suivant le protocole déjà établi de Banerjee *et al.* 2004. L'expression de la protéine s'effectue dans la souche *Escherichia coli* BL21 (DE3) Gold pLysS. Pour purifier la protéine non taguée, une colonne d'affinité est réalisée utilisant un inhibiteur connu de la HCA-II, le pAMBS (p-aminométhylbenzène sulfonamide).

# 3. Biophysique

## 3.1. Dichroïsme circulaire induit (ICD)

Le dichroïsme circulaire repose sur la capacité d'un échantillon possédant un chromophore chiral, ou placé dans un environnement asymétrique, d'absorber différemment la lumière polarisée circulairement droite et la lumière polarisée circulairement gauche (**Figure V.3.**). Le spectre dichroïque correspond à la différence d'absorbance entre ces deux types de lumière, pour chaque longueur d'onde.



**Figure V.3. : Schéma explicatif du phénomène de dichroïsme circulaire.**

Deux ondes circulairement polarisées droite et gauche (vert et rouge) sont absorbées de manière différente par un échantillon optiquement actif, matérialisé par la cuve orange. L'onde résultante (bleu clair) n'est plus verticale mais subie une rotation d'un angle  $\theta$ .

Les foldamères synthétisés sont des entités racémiques. Il y a en solution autant de molécules adoptant une conformation d'hélice droite que de molécules adoptant une conformation d'hélice gauche. Ces molécules sont en outre achirales, puisqu'elles ne possèdent pas de carbone asymétrique. En présence d'une molécule chirale comme une protéine (rappelons que le C $\alpha$  de chaque acide aminé est asymétrique), la distribution hélice droite / hélice gauche peut changer. Cette modification induit l'apparition d'un signal de dichroïsme circulaire autour de 400 nm d'autant plus marqué qu'un type d'hélice est plus présent que l'autre.

Les expériences de dichroïsme circulaire ont été réalisées sur un spectromètre JASCO J-815, sous un flux constant d'azote, permettant de réaliser des mesures d'absorbance UV de 190 à 1100 nm. L'accumulation de deux scans consécutifs avec un défilement adaptatif (60 nm/min par défaut) et un incrément de 1 nm, conduit au spectre. Chaque spectre est corrigé par la soustraction d'un spectre de référence, celui du tampon d'analyse utilisé contenant la protéine seule. Les résultats sont exprimés en extinction molaire par résidu quinoline  $\epsilon_\lambda$  (L.mol<sup>-1</sup>.cm<sup>-1</sup>.res<sup>-1</sup>), basés sur la masse moléculaire du foldamère du composé utilisé. L'ellipticité molaire par résidu est déterminée telle que :

$$[\theta]_\lambda = \frac{\theta_\lambda \times MW}{32\,980 \times C \times l \times n}$$

Avec  $C$  : la concentration du foldamère exprimée en mg/mL.

$l$  : la longueur du chemin optique exprimée en cm.

$MW$  : la masse moléculaire du foldamère en g/mol.

$\theta_\lambda$  : l'ellipticité mesurée en millidegrés à une longueur d'onde donnée.

$n$  : le nombre d'unité quinoline du foldamère.

Toutes les mesures ont été réalisées dans un tampon NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 7,5, avec une concentration en protéine de 1 mg/mL et 1 équivalent de composé sauf mention contraire. Les spectres sont enregistrés à différents temps à 25 °C.

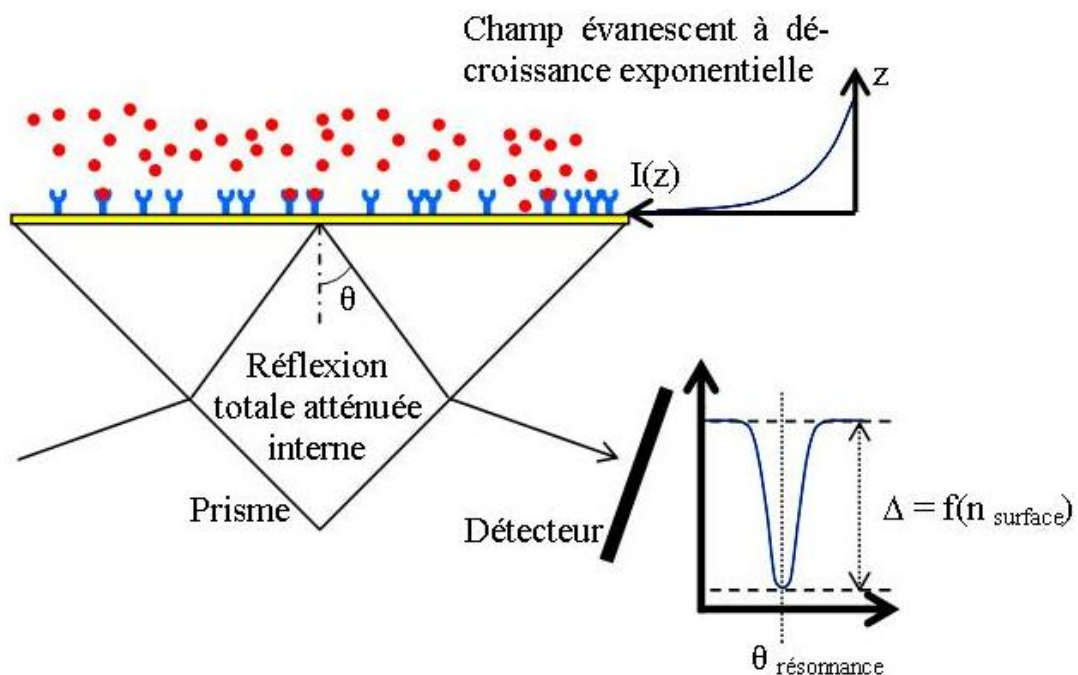
## 3.2. Résonance plasmonique de surface (SPR)

Un système de détection SPR permet de mesurer, en biologie, les constantes d'association et de dissociation de complexes moléculaires.

Les plasmons sont des ondes oscillantes de densité surfacique qui se déplacent parallèlement aux interfaces métal (or, argent) / milieux biologiques (ou air). L'amplitude du

champ électromagnétique des plasmons dans le milieu biologique décroît exponentiellement quand on s'éloigne de la surface, c'est l'onde évanescente de surface. Ces plasmons peuvent être excités par de la lumière dans des conditions bien particulières d'épaisseur de la couche de métal et de réflexion totale atténuée interne. Pour cela le film métallique est posé sur un système optique permettant d'être dans ces conditions. Le plasmon va absorber par résonance une fraction de l'énergie du faisceau incident. Il se forme ainsi un minimum dans l'intensité réfléchie en fonction de l'angle d'incidence  $\theta$  (**Figure V.4.**). Cet angle dépend très fortement du profil de l'indice de réfraction dans l'épaisseur du champ évanescent. Le changement de composition (tampons différents, adsorption, désorption de ligands) au niveau de l'interface peut potentiellement changer cet indice et de ce fait changer l'angle de résonance. Le suivi de la variation de l'angle de résonance en fonction du temps permet de tracer une cinétique d'association et une cinétique de dissociation. A partir des constantes d'association  $k_{on}$  et de dissociation  $k_{off}$ , il est possible de calculer la constante d'affinité  $K_D$  :

$$K_D = \frac{k_{off}}{k_{on}}$$



**Figure V.4. : Fonctionnement d'une expérience de SPR.**

Le changement d'environnement à proximité de la surface du biocapteur va changer l'indice de réfraction  $n_{\text{surface}}$  du milieu et perturber le plasmon. Cette perturbation sera visualisée sur le détecteur et pourra après analyse du signal donner un sensorgramme.

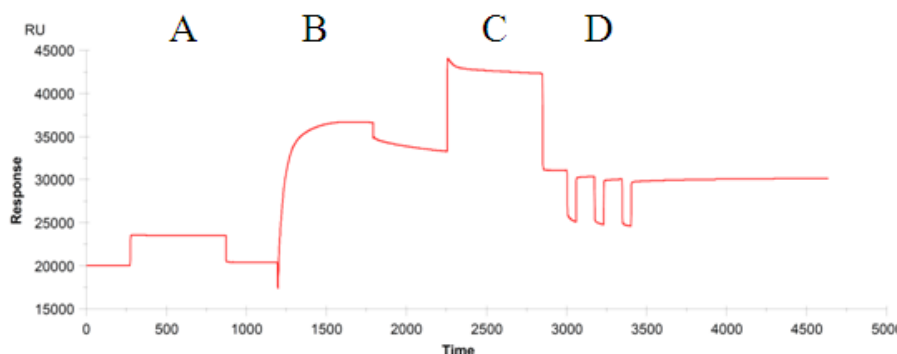
Les mesures ont été réalisées sur un appareil BIAcore 3000 (Biacore, GE Healthcare, Uppsala, Suède), à 25 et/ou 35 °C. Les biocapteurs, le NHS, l'EDC et l'éthanolamine sont commercialisés par GE Healthcare. La HCA-II et le composé contrôle, le furosémide, proviennent de chez Sigma-Aldrich. Tous les tampons et solutions utilisés ont été systématiquement filtrés et dégazés avant utilisation à l'exception des solutions de foldamères qui ont été purifiées par HPLC.

Dans un premier temps, il est important de fixer un des partenaires de l'interaction mesurée à la surface du biocapteur. Il existe plusieurs approches d'immobilisation incluant l'immobilisation covalente par méthodes chimiques, la capture de haute affinité ou encore l'adsorption hydrophobique de mono- ou de bicouches lipidiques. Les protéines, et en particulier la HCA-II, se lient facilement de manière covalente aux matrices de dextran utilisés comme biocapteur (type CM5).

- La première étape consiste à activer les groupements carboxyles libres présents à la surface dextran du biocapteur par injection pendant 10 min à une vitesse de 10  $\mu\text{L}/\text{min}$  d'agents de couplage peptidique, NHS et EDC, à une concentration respective de 0,2 M et 0,05 M (**Figure V.5.**).
- La solution de protéine, préalablement préparée à une concentration de 0,3 mg/mL dans un tampon acétate de sodium 10 mM pH 6,4, est injectée pendant 10 min à la même vitesse. Les résidus lysines réagissent alors avec les esters activés de la matrice dextran pour former une liaison amide covalente.
- L'injection pendant 10 min à la même vitesse d'une solution d'éthanolamine à 1 M pH 8,5, un composé présentant une fonction amine libre, permet de désactiver les groupements esters succinimides n'ayant pas réagi avec la protéine, évitant ainsi toute fixation covalente par les fonctions amines des ligands testés.
- La dernière étape consiste à nettoyer le biocapteur avec trois injections successives d'hydroxyde de soude 10 mM, permettant de décrocher toute molécule non liée covalamment et de stabiliser le signal. Cette étape n'est toutefois pas anodine pour la protéine. Elle est en effet susceptible de dénaturer partiellement ou complètement la protéine fixée. Aussi avons-nous décidé d'omettre cette étape de « lavage » du biocapteur avant utilisation. En contre partie, nous laissons tout de même le signal se stabiliser durant 5 à 10 min avant toute mesure.

Tous les sensorgrammes sont doublement corrigés (Myszka 1999). Il est nécessaire d'une part de soustraire le signal du tampon qui est mesuré sur la même piste avant d'injecter

le ligand à tester. Il est également important de corriger le sensorgramme par le signal que produit le ligand sur une piste ne portant pas la protéine.



**Figure V.5. : Sensorgramme obtenu pendant l'immobilisation d'une protéine sur le biocapteur.**

A correspond à l'étape d'activation par le mélange NHS/EDC. L'immobilisation de la protéine est en B et le blocage des fonctions n'ayant pas réagi avec l'éthanolamine est en C. Finalement l'étape de lavage par trois injections de soude est en D mais n'est jamais réalisée, car cette étape semble dénaturer la protéine. Ce sensorgramme a été obtenu après double correction.

Une fois la protéine immobilisée, trois injections de concentration croissante en furosémide sont réalisées afin de vérifier que la protéine fixée est active. Ce composé est un inhibiteur connu de la HCA-II dont les paramètres cinétiques sont clairement définis ( $K_d \approx 1 \mu\text{M}$ ) (Papalia *et al.* 2006).

La mesure des paramètres cinétiques pour un composé est réalisée grâce à trois injections successives et de concentration croissante du ligand voulu (ces valeurs sont indiquées sur les sensorgrammes). Pendant la dissociation seul le tampon est injecté. Les données sont ensuite analysées par le logiciel BiaEvaluation avec un modèle de 1:1 (Palau et DiPrimo 2013).

## 4. Biologie structurale

### 4.1. Modélisation moléculaire

Les travaux de modélisation moléculaire ont été conduits par le Dr. Michel Laguerre au sein de notre laboratoire. Afin d'étudier les interactions entre une protéine et un foldamère

*in silico*, deux champs de forces ont été utilisés : CHARMM General Force Field (CgenFF) pour le foldamère et CHARMM-22/CMAP pour la protéine. Ces travaux avaient pour but, dans le contexte de synthèse des composés en phase liquide, de préciser la nature et la longueur de l'espaceur devant être intercalé entre l'inhibiteur et le foldamère. Cet espaceur doit donner au foldamère de nombreux degrés de liberté lui permettant d'explorer une large zone et d'interagir avec la surface protéique.

## 4.2. Diffraction des rayons X

La résolution d'une structure 3D de protéine par cristallographie nécessite trois étapes indispensables. La première consiste en l'obtention d'un échantillon d'une très grande pureté et d'une homogénéité importante. Les étapes de production et de purification des deux protéines sur lesquelles j'ai travaillé sont détaillées dans le Chapitre IV. paragraphe 2.3. à 2.6. La seconde étape est l'obtention d'un monocristal exploitable, et enfin la dernière consiste en la collecte du jeu de diffraction et le traitement des données jusqu'à obtenir la structure tridimensionnelle.

### 4.2.1. Cristallogenèse

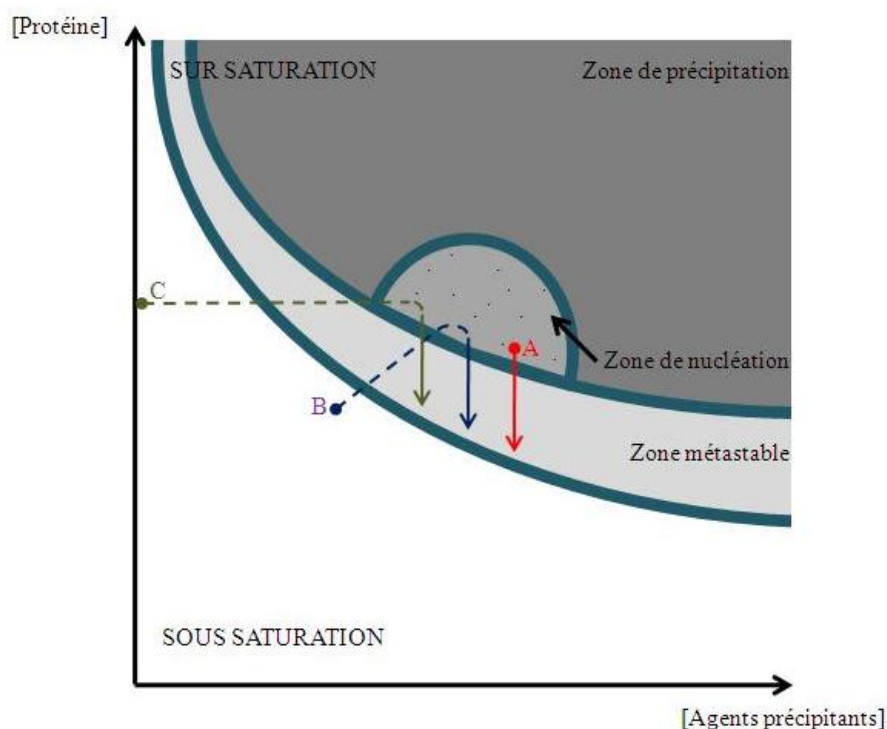
L'ensemble du travail de cristallographie est entièrement conditionné par l'obtention de cristaux de qualité suffisante.

La recherche des premières conditions de cristallisation, puis leur optimisation sont aujourd'hui encore menées de manière très empirique. Malgré cela, il existe de nombreux outils utiles pour nous aider à obtenir des cristaux exploitables.

#### 4.2.1.1. Principe

Selon l'Union Internationale de Cristallographie (IUCr) « Tout solide dont le diffractogramme est essentiellement discret est un cristal ». Le principe de la cristallogenèse est d'induire, à partir d'une protéine initialement en solution, la formation d'un cristal c'est-à-dire un solide dans lequel les molécules s'ordonnent dans les trois directions de l'espace. Ce dernier apparaît (nucléation) et se développe (croissance cristalline) au sein de la phase liquide. Le défi consiste à trouver le compromis entre les différents facteurs cinétiques et thermodynamiques qui contrôlent la solubilité de la protéine. Pour induire la nucléation, il est nécessaire d'amener la protéine dans un état de sursaturation en diminuant progressivement sa solubilité. Ceci revient à déstabiliser les interactions protéines - solvant au profit des

interactions protéines - protéines. Des interactions aléatoires conduisent à la formation de précipité (précipité amorphe), contrairement à la répétition d'interactions particulières qui elles conduisent à la formation de germes cristallins.



**Figure V.6. : Diagramme de phase et évolution du degré de sursaturation en fonction de la technique.**  
**En A la méthode dite en « micro batch », en B la diffusion de vapeur, et en C la diffusion liquide-liquide.**

Les paramètres physico-chimiques qui influencent la solubilité des protéines sont très nombreux et variables d'une protéine à l'autre.

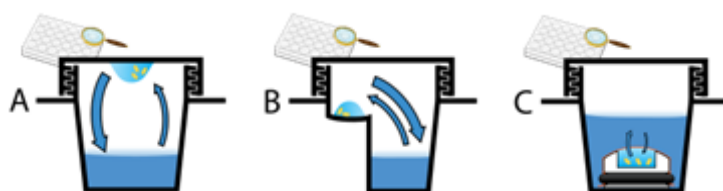
- La concentration de la protéine, sa stabilité en solution, sa pureté (absence ou non de contaminants), l'homogénéité de la solution (présence d'un ou plusieurs états conformationnels) sont parmi les premiers facteurs affectant le processus de cristallogenèse.
- La température, le pH, la force ionique (due à la présence de sels en solution) affectent la cinétique de cristallisation et sont de ce fait des paramètres importants de la cristallogenèse.
- La présence de molécules de substrat, d'inhibiteur, ou de cofacteurs sont aussi des facteurs importants. La nature et la concentration de divers agents précipitants et/ou d'additifs sont également décisives pendant la cristallogenèse.



#### 4.2.1.2. Méthodes et techniques

Il existe de nombreuses méthodes de cristallisation des protéines, chacune permettant, par des chemins différents, d'atteindre les conditions de nucléation. La diffusion de vapeur, la diffusion liquide-liquide ou encore l'évaporation sont les méthodes les plus utilisées pour obtenir un cristal. Chacune de ses méthodes peut être adaptée en différentes techniques.

La méthode de diffusion de vapeur est la plus couramment utilisée (McPherson, 1982). Elle est basée sur les lois thermodynamiques qui imposent aux concentrations en soluté de deux solutions, enfermées dans un conteneur hermétiquement clos, de tendre vers un équilibre par diffusion des composés les plus volatils. En début d'expérience, une goutte de protéine est diluée deux fois par une solution réservoir contenant un tampon, un ou des agents précipitants et/ou des sels. Cette goutte est placée dans une enceinte hermétiquement scellée (par de la graisse silicone), en présence d'un plus important volume de solution réservoir. La concentration en sel et/ou en agents précipitants étant plus élevée dans la solution réservoir que dans la goutte, le retour à l'équilibre des concentrations est atteint essentiellement par diffusion des molécules d'eau de la goutte vers le réservoir. La diminution du volume de la goutte entraîne l'augmentation des concentrations des agents précipitants et de la protéine et donc le « passage » dans la zone de nucléation (**Figure V.6.B**). Différentes techniques (goutte suspendue ou assise) existent, toutefois le principe exposé ci-dessus reste le même (**Figure V.7.A, B**).



**Figure V.7. Différentes techniques en diffusion de vapeur et liquide-liquide.**

**A :** La configuration en goutte suspendue. La technique de la goutte assise est représentée en **B** et l'utilisation du bouton de dialyse est schématisée en **C**.

La technique de dialyse fait appel au principe de diffusion liquide-liquide. Elle utilise un petit volume du mélange protéine-solution précipitante placé dans un « bouton de dialyse » lui-même disposé dans un volume plus important de solution précipitante (**Figure V.7.C**). Seuls les agents précipitants ont la capacité de diffuser librement à travers la membrane semi-

perméable de dialyse. Pour atteindre l'équilibre du système, la concentration en agents précipitants dans le bouton de dialyse va augmenter, faisant passer la protéine dans un état de sursaturation. La formation de germes cristallins puis de cristaux fera à nouveau diminuer la concentration en protéine (**Figure 6.C.**).

La méthode d'évaporation en batch est utilisée lorsque le mélange protéine-solution précipitante est placé ou non sous une couche d'huile plus ou moins perméable à l'eau. L'évaporation contrôlée de l'eau permet de concentrer la protéine et de l'amener doucement dans un état de sur saturation propice à la formation de germe.

Le micro-batch est relativement différent des méthodes précédentes puisqu'une fois l'expérience réalisée, plus aucun changement n'intervient. Cette méthode nécessite de mélanger une solution protéique avec une solution précipitante et de placer ce mélange sous une couche d'huile imperméable. Le mélange doit être réalisé de telle manière que la protéine soit immédiatement dans un état de sursaturation propice à l'apparition de nucléi (**Figure V.6.A**). Le système, dans ce cas, ne subira aucune diffusion ou évaporation, ici seuls la nature et les concentrations des précipitants, des solvants et de la protéine sont importantes.

Lorsque seuls des cristaux de trop petite taille sont obtenus, l'ensemencement par *streaking* (ou *micro-seeding*) peut être d'une très grande aide. Le *streaking* consiste à promouvoir la nucléation par introduction de germes dans une solution sursaturée. Cela se fait généralement en tirant un trait dans une goutte de cristallisation pré-équilibrée à l'aide d'un poil de moustache de chat (ou de tout autre objet fin et poreux) chargé préalablement en germes par contact préalable avec des microcristaux. Il est également possible d'introduire dans une goutte pré-équilibrée non plus des germes mais de petits cristaux préalablement lavés dans une solution stabilisée. On parle alors de *macro-seeding*. La méthode de l'ensemencement permet théoriquement de contrôler le nombre de cristaux dans la goutte de cristallisation et donc leur taille. Par ailleurs en dissociant les étapes de nucléation et de croissance, elle permet de faire pousser les cristaux dans des conditions moins saturantes. Les cristaux poussent souvent plus lentement et donc plus régulièrement dans des conditions de faible saturation. Ces techniques peuvent paraître empiriques mais donnent de très bons résultats.

#### 4.2.1.3. Recherche des conditions et optimisation

La totalité des paramètres physico-chimiques qui peuvent influencer la cristallisation des protéines définit un espace paramétrique important propre à chaque protéine ou complexe

macromoléculaire. La quantité d'échantillon est très souvent limitée, aussi est-il nécessaire de parcourir cet espace en un minimum d'expériences, on parle de criblage initial des conditions de cristallisation ou *screening* en anglais.

Le criblage consiste donc à réaliser le minimum d'expériences tout en couvrant le plus possible l'espace paramétrique. Pour cela des matrices de criblage ont été mises au point (elles peuvent contenir les conditions de cristallisation ayant démontré une réelle efficacité pour d'autres protéines) tout en permettant d'évaluer une large gamme de pH, d'agents précipitants et d'additifs. Les kits utilisés en routine au laboratoire sont les Crystal Screen I et II, Salt RX ainsi que Index commercialisés par Hampton Research, et JCSG+ ou encore Morpheus commercialisés par Molecular Dimensions. Il existe de très nombreux autres kits de criblage. Pour réaliser ces expériences, des plaques CrystalQuick 96 puits (Greiner Bio One, Allemagne) sont utilisées contenant 40 à 80  $\mu\text{L}$  de solution de réservoir. Les gouttes assises sont composées de 400 nL à 1  $\mu\text{L}$  de la solution du réservoir et le même volume de la solution de protéine. Ainsi le criblage permet de dégager rapidement, sans connaissance préalable, les premières tendances à la cristallisation d'une protéine ou d'un complexe, en utilisant peu de matériel.

Un certain nombre de conditions de cristallisation peuvent faire apparaître des cristaux. Ils sont généralement petits et souvent grossiers et leur diffraction est fréquemment de piètre qualité. Aussi est-il utile de réaliser des expériences d'optimisation pour obtenir des cristaux plus gros, plus réguliers et de meilleure qualité diffractante. Le but de cette étape est d'ajuster finement les conditions afin de diminuer le nombre de germes, d'augmenter la taille des cristaux, de minimiser les défauts cristallins (macles, présence d'oursins, cristaux accolés, ...) afin d'augmenter l'ordre cristallin.

Les expériences d'optimisation exécutées au cours de ces travaux, ont été réalisées dans des plaques Linbro 24 puits avec 500  $\mu\text{L}$  de solution de réservoir et des gouttes suspendues de 2  $\mu\text{L}$  (1  $\mu\text{L}$  de la solution protéique et 1  $\mu\text{L}$  de la solution du réservoir). Tous les paramètres d'une condition issue du criblage pourront être affinés, mais en général le pH, et les concentrations en précipitants ou en additifs sont les premiers paramètres qui sont ajustés en plus de la température. Deux paramètres peuvent être affinés à la fois par plaque suivant les deux directions horizontale et verticale de la plaque.

#### 4.2.1.4. Congélation des cristaux

Lors des expériences de diffraction, les cristaux sont exposés plus ou moins longtemps aux rayons X. Ces rayonnements engendrent l'apparition de radicaux libres qui se propagent (Massover 2007) et provoquent la dégradation des cristaux (détérioration progressive des ponts disulfures, décarboxylation des aspartates et des glutamates, perte des groupements OH des tyrosines, ...). Avec l'emploi de sources de rayon X de plus en plus intenses, la nécessité de protéger les cristaux contre ces phénomènes de dégradation se fait de plus en plus ressentir. La congélation des cristaux et la diffraction à très basse température (100 °K) permet, à défaut, de diminuer l'apparition de ces radicaux libre, et leur propagation dans le cristal.

La congélation des cristaux se fait par trempage dans de l'azote liquide, après l'ajout d'un cryoprotectant. Les cristaux de macromolécules biologiques contiennent énormément d'eau (de 30 à 70 %) ce qui peut entraîner la formation de glace cristalline lors de la congélation. La formation de cristaux de glace, en plus de faire apparaître des signaux parasites lors de la diffraction (principaux anneaux de glace à 2,2 ; 2,6 ; 3,4 ; 3,6 ; 3,8 Å de résolution), peut faire augmenter le volume du cristal, le détériorer, voire même le détruire. Pour éviter ce problème, une méthode consiste à tremper les cristaux dans des solutions de réservoirs complémentées avec des concentrations croissantes en cryoprotectant dans le but d'échanger, sans trop de perturbation pour le cristal, une partie des molécules d'eau par des molécules de cryoprotectant. Divers cryoprotectants sont habituellement utilisés au laboratoire tels que le glycérol, le MPD, le PEG 400, l'huile de Paratone à une concentration finale pouvant atteindre 30 % v/v ou w/v environ.

#### 4.2.2. Cristallographie

##### 4.2.2.1. Diffraction et traitement des données

L'obtention de la structure d'une macromolécule biologique, nécessite d'observer des distances de l'ordre de l'Angström ( $1 \text{ Å} = 0,1 \text{ nm}$ ). « Voir » à cette résolution demande l'utilisation d'un rayonnement électromagnétique dont la longueur d'onde doit être du même ordre de grandeur, c'est-à-dire de l'Angström, gamme de longueurs d'onde à laquelle appartiennent les rayons X, **Figure V.8**.

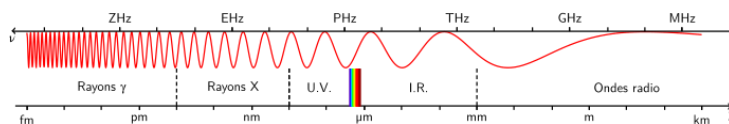


Figure V.8. : Spectre électromagnétique.

L'axe du haut correspond à la fréquence de l'onde et l'axe du bas à sa longueur d'onde. Les RX ont des longueurs d'onde allant d'environ 10 nm à 10 pm.

C'est le motif cristallin qui permet le phénomène de diffraction, encore appelé diffusion élastique des RX par les électrons des atomes qui composent le cristal. Ces ondes diffusent dans toutes les directions de l'espace. Elles n'ont cependant pas toutes la même phase lorsqu'elles frappent les atomes (déphasage rouge dans la **Figure V.9.**). De plus, en un point de l'espace, les ondes diffusées proviennent de tous les atomes du cristal et subissent encore un déphasage dû à la différence de chemin optique (déphasage vert dans la **Figure V.9.**). Dans certains endroits de l'espace, les ondes s'annulent (interférences destructives), et dans d'autres, les ondes s'additionnent et l'on mesure une intensité positive. C'est la distance inter-réticulaire  $d$  du plan cristallin diffractant, qui sera le discriminant entre une interférence destructive ou constructive. C'est la loi de Bragg :

$$2 \cdot d \sin \theta = n \cdot \lambda$$

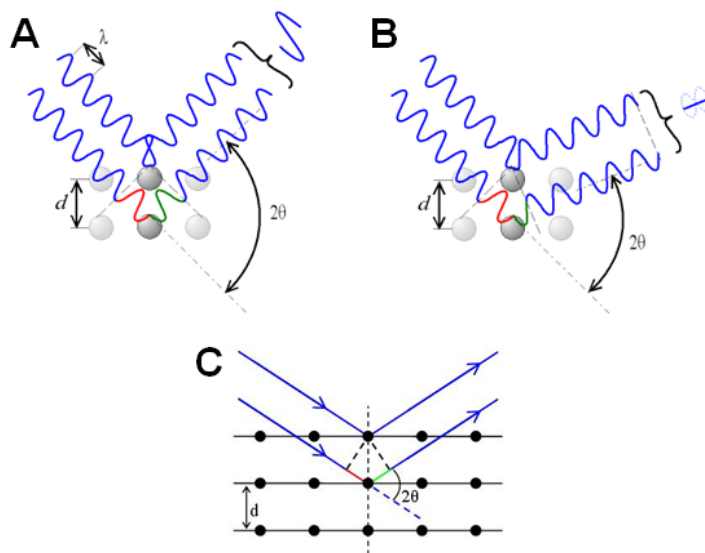


Figure V.9. : Illustration de la loi de Bragg.

Pour qu'il y ait interférence constructive comme en A la somme du déphasage du faisceau incident (en rouge ( $d \sin \theta$ )) et du déphasage du faisceau diffusé (en vert ( $d \sin \theta$ )) doit être égale à un multiple de la longueur d'onde  $\lambda$ . Si cette égalité n'est pas respectée, il y a déphasage comme en B et aucune onde diffractée ne sera mesurée. En C le schéma est simplifié.

Le traitement des données de l'expérience de diffraction peut être réalisé par un très grand nombre de programmes tels que MOSFLM (Leslie 2006) ou encore XDS (Kabsch 2010). Ces deux programmes sont utilisés pour l'indexation des intensités diffractées. Ils proposent, à partir de quelques clichés de diffraction, une liste de types de réseau, les paramètres de maille et les groupes d'espace potentiels ainsi que l'orientation du cristal, et permettent une évaluation semi-automatique des résultats. Une fois le groupe d'espace et les paramètres de maille choisis, ils intègrent les intensités diffractées. Les intensités mesurées doivent être corrigées des effets de la polarisation du faisceau et des effets d'irradiation. La mise à l'échelle s'effectue principalement avec les programmes SCALA (Evans 2006), de la suite de programmes CCP4 (Winn *et al.* 2011) ou avec XDS.

Le rapport signal sur bruit  $I/\sigma(I)$  permet de définir la limite de résolution du jeu de données. Seules les intensités ayant une résolution supérieure à cette limite sont conservées. Ce paramètre  $I/\sigma(I)$ , dépend du cristal mais aussi des conditions d'enregistrement : un temps de pose élevé et donc un nombre élevé de coups  $N_c$  par réflexion, améliore la précision statistique des données (le rapport signal sur bruit est approximativement proportionnel à  $\sqrt{N_c}$ ).

Pour juger de la qualité d'un jeu de données, le rapport signal sur bruit est très important, mais la redondance (nombre de mesures d'une même intensité) ainsi que la complétude (nombre d'intensités mesurées sur le nombre d'intensités mesurables) le sont également. Deux autres facteurs résiduels ( $R_{merge}$  et  $R_{meas}$ ) doivent également être considérés. Ils indiquent l'accord entre les intensités équivalentes par symétrie et leur moyenne en tenant compte ou non de la redondance.

$$R_{merge} = \frac{\sum_{hkl} \sum_i^N |I_{hkl} - \bar{I}_{hkl}^i|}{\sum_{hkl} \sum_i^N |I_{hkl}^i|}$$

$$R_{meas} = \frac{\sum_{hkl} \left(\frac{N}{N-1}\right)^{1/2} \sum_i^N |I_{hkl} - \bar{I}_{hkl}^i|}{\sum_{hkl} \sum_i^N |I_{hkl}^i|}$$

A la fin du traitement des données, une liste d'intensités diffractées  $I(hkl)$  est obtenue.

Les intensités  $I(hkl)$  sont proportionnelles au carré du module des facteurs de structure  $\overrightarrow{F_{hkl}}$ . Les facteurs de structure sont des nombres complexes avec une amplitude  $\|\overrightarrow{F_{hkl}}\|$  et une

phase  $\varphi_{hkl}$  que l'on peut écrire  $\overrightarrow{F_{hkl}} = \|\overrightarrow{F_{hkl}}\| e^{i\varphi_{hkl}}$ . La mesure de l'intensité  $I(hkl)$  ne peut fournir que le module du facteur de structure des atomes. La phase  $\varphi_{hkl}$  est indispensable pour le calcul de la carte de densité électronique, mais ne peut être obtenue expérimentalement à partir d'un seul jeu de données enregistré avec la protéine native. C'est ce que l'on appelle communément « le problème des phases »

#### 4.2.2.2. Résolution du problème des phases

Il existe plusieurs méthodes pour résoudre au mieux ce problème. Toutes, à l'exception d'une seule (celle dite du remplacement moléculaire), sont basées sur le remplacement isomorphe et/ou la mesure d'un signal anomal (SIR, SIRAS, MIR, MIRAS, et MAD). Au cours de nos travaux, nous avons exclusivement utilisé la méthode du remplacement moléculaire. Nous n'introduirons donc que très brièvement les autres méthodes.

##### 4.2.2.2.1. Remplacement isomorphe

La méthode du remplacement isomorphe consiste à insérer un ou plusieurs atomes lourds (atomes possédant un grand nombre d'électrons comparativement à ceux, C, N, O, S, généralement présents dans les protéines) dans l'édifice cristallin de la protéine native. Le problème des phases est ensuite résolu à partir de la contribution de ces seuls atomes lourds. Pour obtenir un dérivé lourd, une étape de trempage, de co-cristallisation ou d'exposition à une atmosphère riche en gaz noble (Xe ou Kr) est nécessaire. Plusieurs dérivés lourds doivent être le plus souvent obtenus. Cette méthode nécessite donc l'obtention et la mesure de plusieurs cristaux (MIR). Elle est d'autant plus lourde à mettre en place qu'elle nécessite en outre un degré d'isomorphisme élevé entre dérivés lourds et protéine native. En d'autres termes, l'introduction de l'atome lourd ne doit que très faiblement perturber l'arrangement cristallin. Grâce à l'utilisation d'une fonction Patterson différence, les positions des atomes lourds peuvent être obtenues et les phases de la protéine native approximées.

La prise en compte du signal anomal du ou des atome(s) lourd(s) (SIRAS, MIRAS) peut singulièrement limiter le nombre de dérivés à cristalliser.

##### 4.2.2.2.2. Diffusion anormale

Dans le cas des diffuseurs anomaux, une contribution imaginaire supplémentaire dans le facteur de diffusion atomique va rendre caduque l'égalité des intensités  $I_{hkl}$  et  $I_{\bar{h}\bar{k}\bar{l}}$  des

paires de Friedel. La prise en compte de la diffusion anormale, et donc la mesure des différences d'intensité entre ces paires, permet de remonter jusqu'aux phases de la protéine native. Cette méthode peut être utilisée en complément du remplacement isomorphe. On parle alors de SIRAS ou de MIRAS. Elle peut également être utilisée seule (MAD) dès lors que différentes collectes à des longueurs d'onde judicieusement choisies sont effectuées. Signalons toutefois que l'utilisation de telles méthodes nécessite l'accès au rayonnement synchrotron (longueur d'onde continûment modulable), une très bonne précision de mesures des intensités et une grande stabilité du cristal sous le faisceau de Rayons X. La méthode MAD est le plus souvent utilisée lorsqu'une protéine est exprimée en présence de sélénométhionine et non de méthionine.

Pour être complet, il nous faudrait parler de SAD. Dans ce cas, il est toutefois nécessaire de faire appel aux méthodes directes (méthodes non expliquées ici).

#### 4.2.2.2.3. Remplacement moléculaire

La méthode dite du remplacement moléculaire tire bénéfice de l'existence de structures 3D de protéines déjà résolues dès lors que l'identité de séquence entre ces protéines dites modèles et la protéine cible (protéine dont on recherche la structure) est supérieure à 30 %. On estime en effet dans ce cas que l'écart quadratique moyen (r.m.s.d.) des positions des Ca est inférieur à 1,5 Å (Clothia 1986) et que les repliements des protéines modèles et cible ont de grandes chances d'être comparables.

En effet la transformée de Fourier des intensités diffractées permet de calculer une carte de Patterson. Cette carte, en l'absence de diffusion anormale, devient la fonction d'auto-corrélation de la densité électronique. Ainsi la position de chaque pic de cette carte correspond au vecteur existant entre deux atomes avec une intensité égale au produit de leur nombre d'électrons. Une telle carte (appliquée aux macromolécules) contient beaucoup trop d'informations et ne permet pas d'obtenir des données structurales de haute résolution. Cependant les cartes de Patterson de la structure à déterminer et de la structure modèle doivent, après correction de leur orientation et de leur position, être très similaires. C'est sur ce principe que repose la méthode de remplacement moléculaire qui est composée de deux étapes : la recherche de la fonction de rotation et de la fonction de translation.

La recherche de la fonction de rotation reliant les deux fonctions de Patterson (celle du modèle et celle de la structure étudiée) est réalisée par des algorithmes utilisant la fonctions de maximum de vraisemblance. Les plus importantes corrélations (scores élevés) sont obtenues



lorsque les deux structures ont la même orientation. La fonction de translation reliant les deux fonctions de Patterson est recherchée de la même manière. Cette méthode est relativement gourmande en temps machine, mais les récentes avancées informatiques et les fonctions de translation et de rotation rapides permettent de réduire les temps d'analyse.

Une fois les deux fonctions de Patterson, et les structures *a fortiori*, correctement alignées, les phases du modèle peuvent être appliquées, en première approximation, à la structure étudiée. Nous utilisons préférentiellement le programme Molrep (Vagin et Teplyakov 1997) de la suite de programme CCP4 pour réaliser ces calculs.

#### 4.2.2.2.4. Amélioration des cartes de densité

Les premières positions atomiques ayant été définies, il devient alors possible de calculer les premières cartes de densité électronique.

$$\rho(x, y, z) = TF^{-1} \overrightarrow{F_{hkl}} = \frac{1}{V} \sum_h \sum_k \sum_l \overrightarrow{F_{hkl}} e^{-2i\pi \times (hx_j + ky_j + lz_j)}$$

De telles cartes sont entachées d'erreurs, les positions atomiques utilisées pour les calculer étant incomplètes et peu précises. Il convient donc de les améliorer.

L'aplatissement de solvant est la première technique à mettre en œuvre lorsqu'une première carte de densité et un premier modèle sont obtenus. Les molécules de solvant comblant l'espace existant entre les chaînes protéiques sont le plus souvent désordonnées. En affectant une densité résiduelle de 0,33 e.Å<sup>-3</sup> à ces zones préalablement déterminées, il est possible d'améliorer les phases.

Dès lors que l'unité asymétrique contient deux ou plusieurs chaînes identiques d'une même protéine, les densités électroniques associées à chaque chaîne doivent être semblables. La moyennisation des densités symétriques non cristallographiques est une seconde technique permettant d'améliorer la première estimation des phases.

#### 4.2.2.3. Affinement et validation du modèle

Une fois le problème des phases résolu et les cartes de densité électronique améliorées, les nouvelles cartes calculées présentent un meilleur contraste. L'interprétation de ces cartes s'effectue avec le logiciel Coot (Emsley 2010). La première étape dans l'élaboration du modèle est de construire la chaîne principale. Les chaînes latérales sont ensuite placées grâce à la connaissance de la séquence primaire si les densités sont visibles dans les cartes. Un tel modèle contient toujours des erreurs et l'affinement consiste à les réduire au maximum en

minimisant la différence entre l'amplitude des facteurs de structure observés expérimentalement  $\|\overrightarrow{F_{obs}}\|$  et l'amplitude des facteurs de structure calculés à partir du modèle  $\|\overrightarrow{F_{calc}}\|$ . Cela revient à minimiser, grâce aux logiciels Refmac5 (Murshudov *et al.* 1997) de la suite CCP4, CNS (Brunger *et al.* 1998) et Phenix (Adams *et al.* 2010), la fonction de maximum de vraisemblance.

Afin de palier au rapport souvent limité entre le nombre d'observations expérimentales (intensités diffractées) et le nombre de paramètres à affiner, la minimisation est faite en tenant compte de contraintes stéréochimiques (longueurs et angles de liaison, angles dièdres, ...) issues de structures à haute résolution d'oligopeptides de poids moléculaire limité. Il existe différentes stratégies permettant de limiter le nombre de paramètres à affiner. Parmi celles-ci, citons :

- L'affinement de type « groupe rigide » qui permet de réduire considérablement le nombre de paramètres à affiner, en considérant une entité moléculaire comme un objet rigide unique localisable dans l'espace par 6 paramètres (3 angles de rotation et 3 vecteurs de translation).
- La prise en compte de facteurs d'agitation thermique isotrope, pouvant être ou non considérés comme identiques pour tous les atomes d'un groupe donné, est une autre façon de limiter le nombre de paramètres.

En fait, le nombre de paramètres pouvant être pris en compte dépend de la résolution à laquelle ont été collectées les données. Meilleure sera la résolution, plus grand pourra être ce nombre. La localisation d'un nombre important de molécules de solvant, la mise en évidence de chaînes alternées nécessitent donc de bonnes données et ce n'est qu'à des résolutions proches de la résolution atomique, que l'agitation thermique d'un atome donné peut être décrite de façon anisotrope.

Deux facteurs résiduels ( $R_{work}$  et  $R_{free}$ ) permettent de suivre la bonne convergence de l'affinement. Le premier,  $R_{work}$ , rend compte de l'adéquation entre le module des facteurs de structure observés (données expérimentales) et le module des facteurs de structure calculés à partir des positions atomiques considérées dans le modèle.  $R_{work}$  est calculé en considérant toutes les intensités diffractées prises en compte dans le modèle, à l'exception de 5 à 10% d'entre elles, choisies aléatoirement et non prises en compte dans l'affinement.

$$R_{work} = \frac{\sum_{hkl} (\|\overrightarrow{F_{obs}}\| - \|\overrightarrow{F_{calc}}\|)}{\sum_{hkl} \|\overrightarrow{F_{obs}}\|}$$

Le second facteur,  $R_{free}$ , est calculé à partir des seules intensités exclues du calcul précédent. Il constitue un indicateur non biaisé de l'avancement de l'affinement et rend compte de la pertinence de toute modification apportée au modèle vis-à-vis de la structure « vraie » recherchée.

$$R_{free} = \frac{\sum_{hkl_{free}} (||\vec{F}_{obs}|| - ||\vec{F}_{calc}||)}{\sum_{hkl_{free}} ||\vec{F}_{obs}||}$$

D'autres critères, en fin d'affinement, peuvent être pris en considération. L'écart-type calculé sur les longueurs et les angles de liaison, ainsi que sur les angles dièdres ; les interactions intermoléculaires non réalistes, les conformations non usuelles des chaînes latérales sont autant de paramètres qui permettent de juger de la vraisemblance du modèle. Le serveur MOLPROBITY (Chen *et al.* 2010) permet, entre autres, de détecter ces problèmes.

Finalement les coordonnées atomiques peuvent être déposées à la Protein Data Bank (PDB) (<http://www.rcsb.org>) permettant une mise à disposition de la structure déterminée auprès de la communauté scientifique.

# Références Bibliographiques

- Abbate, F., Supuran, C.T., Scozzafava, A., Orioli, P., Stubbs, M.T., and Klebe, G. (2002). Nonaromatic sulfonamide group as an ideal anchor for potent human carbonic anhydrase inhibitors: role of hydrogen-bonding networks in ligand binding and drug design. **J. Med. Chem.** 45, 3583–3587.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallogr. D Biol. Crystallogr.** 66, 213–221.
- Alberts, I.L., Nadassy, K., and Wodak, S.J. (1998). Analysis of zinc binding sites in protein crystal structures. **Protein Sci.** 7, 1700–1716.
- Appella, D.H., Christianson, L.A., Klein, D.A., Powell, D.R., Huang, X., Barchi, J.J., Jr, and Gellman, S.H. (1997). Residue-based control of helix shape in beta-peptide oligomers. **Nature** 387, 381–384.
- Arvidsson, P.I., Ryder, N.S., Weiss, H.M., Hook, D.F., Escalante, J., and Seebach, D. (2005). Exploring the antibacterial and hemolytic activity of shorter- and longer-chain beta-, alpha,beta-, and gamma-peptides, and of beta-peptides from beta2-3-aza- and beta3-2-methylidene-amino acids bearing proteinogenic side chains--a survey. **Chem. Biodivers.** 2, 401–420.
- Aster J., Kumar V., Abbas A.K., Fausto N. (2009) **Robbins & Cotran Pathologic Basis of Disease**. Philadelphia, Saunders. 54
- Avvaru, B.S., Kim, C.U., Sippel, K.H., Gruner, S.M., Agbandje-McKenna, M., Silverman, D.N., and McKenna, R. (2010). A short, strong hydrogen bond in the active site of human carbonic anhydrase II. **Biochemistry** 49, 249–251.
- Bader, G.D., Betel, D., and Hogue, C.W.V. (2003). BIND: the Biomolecular Interaction Network Database. **Nucleic Acids Res.** 31, 248–250.
- Banerjee, A.L., Swanson, M., Roy, B.C., Jia, X., Haldar, M.K., Mallik, S., and Srivastava, D.K. (2004). Protein surface-assisted enhancement in the binding affinity of an inhibitor for recombinant human carbonic anhydrase-II. **J. Am. Chem. Soc.** 126, 10875–10883.
- Baptiste, B., Douat-Casassus, C., Laxmi-Reddy, K., Godde, F., and Huc, I. (2010). Solid phase synthesis of aromatic oligoamides: application to helical water-soluble foldamers. **J. Org. Chem.** 75, 7175–7185.
- Borish, L.C., Nelson, H.S., Corren, J., Bensch, G., Busse, W.W., Whitmore, J.B., Agosti, J.M., and IL-4R Asthma Study Group (2001). Efficacy of soluble IL-4 receptor for the treatment of adults with asthma. **J. Allergy Clin. Immunol.** 107, 963–970.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. **Acta Crystallogr. D Biol. Crystallogr.** 54, 905–921.

- Chames, P., Van Regenmortel, M., Weiss, E., and Baty, D. (2009). Therapeutic antibodies: successes, limitations and hopes for the future. **Br. J. Pharmacol.** 157, 220–233.
- Chen, E., Paing, M.M., Salinas, N., Sim, B.K.L., and Tolia, N.H. (2013). Structural and functional basis for inhibition of erythrocyte invasion by antibodies that target *Plasmodium falciparum* EBA-175. **PLoS Pathog.** 9, e1003390.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallogr. D Biol. Crystallogr.** 66, 12–21.
- Cheng, R.P., Gellman, S.H., and DeGrado, W.F. (2001). beta-Peptides: from structure to function. **Chem. Rev.** 101, 3219–3232.
- Clackson, T., and Wells, J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. **Science** 267, 383–386.
- Chothia, C., and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. **EMBO J.** 5, 823–826.
- Dawson, C.H., Brown, B.L., and Dobson, P.R. (1997). A 70-kDa protein facilitates interleukin-4 signal transduction in the absence of the common gamma receptor chain. **Biochem. Biophys. Res. Commun.** 233, 279–282.
- DeLano, W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. **Curr. Opin. Struct. Biol.** 12, 14–20.
- Delsuc, N., Kawanami, T., Lefeuvre, J., Shundo, A., Ihara, H., Takafuji, M., and Huc, I. (2008). Kinetics of helix-handedness inversion: folding and unfolding in aromatic amide oligomers. **Chemphyschem** 9, 1882–1890.
- Demchuk, E., Mueller, T., Oschkinat, H., Sebald, W., and Wade, R.C. (1994). Receptor binding properties of four-helix-bundle growth factors deduced from electrostatic analysis. **Protein Sci.** 3, 920–935.
- Disney, M.D., Hook, D.F., Namoto, K., Seeberger, P.H., and Seebach, D. (2005). N-linked glycosylated beta-peptides are resistant to degradation by glycoamidase A. **Chem. Biodivers.** 2, 1624–1634.
- Domingues, H., Cregut, D., Sebald, W., Oschkinat, H., and Serrano, L. (1999). Rational design of a GCN4-derived mimetic of interleukin-4. **Nat. Struct. Biol.** 6, 652–656.
- Duppatla, V., Gjorgjevikj, M., Schmitz, W., Kottmair, M., Mueller, T.D., and Sebald, W. (2012). Enzymatic deglutathionylation to generate interleukin-4 cysteine muteins with free thiol. **Bioconjug. Chem.** 23, 1396–1405.
- Edgington, P.R., McCabe, P., Macrae, C.F., Pidcock, E., Shields, G.P., Taylor, R., Towler, M., and Van De Streek, J. (2006). Mercury: visualization and analysis of crystal structures. **Journal of Applied Crystallography** 39, 453–457.
- Ellman, G.L. (1958). A colorimetric method for determining low concentrations of mercaptans. **Arch. Biochem. Biophys.** 74, 443–450.

- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. **Acta Crystallogr. D Biol. Crystallogr.** 66, 486–501.
- Erlanson, D.A., Wells, J.A., and Braisted, A.C. (2004). Tethering: fragment-based drug discovery. **Annu Rev Biophys Biomol Struct** 33, 199–223.
- Evans, P. (2006). Scaling and assessment of data quality. **Acta Crystallogr. D Biol. Crystallogr.** 62, 72–82.
- Fazal, M.A., Roy, B.C., Sun, S., Mallik, S., and Rodgers, K.R. (2001). Surface recognition of a protein using designed transition metal complexes. **J. Am. Chem. Soc.** 123, 6283–6290.
- Ferrand, Y., Kendhale, A.M., Kauffmann, B., Grélard, A., Marie, C., Blot, V., Pipelier, M., Dubreuil, D., and Huc, I. (2010). Diastereoselective encapsulation of tartaric acid by a helical aromatic oligoamide. **J. Am. Chem. Soc.** 132, 7858–7859.
- Finkelman, F.D., Madden, K.B., Morris, S.C., Holmes, J.M., Boiani, N., Katona, I.M., and Maliszewski, C.R. (1993). Anti-cytokine antibodies as carrier proteins. Prolongation of in vivo effects of exogenous cytokines by injection of cytokine-anti-cytokine antibody complexes. **J. Immunol.** 151, 1235–1244.
- Gadek, T.R., Burdick, D.J., McDowell, R.S., Stanley, M.S., Marsters, J.C., Jr, Paris, K.J., Oare, D.A., Reynolds, M.E., Ladner, C., Zioncheck, K.A., et al. (2002). Generation of an LFA-1 antagonist by the transfer of the ICAM-1 immunoregulatory epitope to a small molecule. **Science** 295, 1086–1089.
- Gellman, S.H. (1998). Foldamers: A Manifesto. **Acc. Chem. Res.** 31, 173–180.
- Gonzalez, M.W., and Kann, M.G. (2012). Chapter 4: Protein interactions and disease. **PLoS Comput. Biol.** 8, e1002819.
- Guenounou, M. (1998). [Cytokines and allergic response]. **Ann. Biol. Clin.** (Paris) 56, 297–304.
- Hage, T., Sebald, W., and Reinemer, P. (1999). Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. **Cell** 97, 271–281.
- Hallett, M.A., Venmar, K.T., and Fingleton, B. (2012). Cytokine stimulation of epithelial cancer cells: the similar and divergent functions of IL-4 and IL-13. **Cancer Res.** 72, 6338–6343.
- Heaney, M.L., and Golde, D.W. (1996). Soluble cytokine receptors. **Blood** 87, 847–857.
- Heck, T., Limbach, M., Geueke, B., Zacharias, M., Gardiner, J., Kohler, H.-P.E., and Seebach, D. (2006). Enzymatic degradation of beta- and mixed alpha,beta-oligopeptides. **Chem. Biodivers.** 3, 1325–1348.
- Hu-Li, J., Shevach, E.M., Mizuguchi, J., Ohara, J., Mosmann, T., and Paul, W.E. (1987). B cell stimulatory factor 1 (interleukin 4) is a potent costimulant for normal resting T lymphocytes. **J. Exp. Med.** 165, 157–172.
- Iriondo-Alberdi, J., Laxmi-Reddy, K., Bouguerne, B., Staedel, C., and Huc, I. (2010). Cellular internalization of water-soluble helical aromatic amide foldamers. **ChemBiochem** 11, 1679–1685.

- Jain, A., Whitesides, G.M., Alexander, R.S., and Christianson, D.W. (1994). Identification of two hydrophobic patches in the active-site cavity of human carbonic anhydrase II by solution-phase and solid-state studies and their use in the development of tight-binding inhibitors. **Journal of Medicinal Chemistry** 37, 2100–2105.
- Janin, J. (1995). Principles of protein-protein recognition from structure to thermodynamics. **Biochimie** 77, 497–505.
- Janin, J., Rodier, F., Chakrabarti, P., and Bahadur, R.P. (2007). Macromolecular recognition in the Protein Data Bank. **Acta Crystallogr. D Biol. Crystallogr.** 63, 1–8.
- Jellinek, D., Green, L.S., Bell, C., and Janjić, N. (1994). Inhibition of receptor binding by high-affinity RNA ligands to vascular endothelial growth factor. **Biochemistry** 33, 10450–10456.
- Jiang, H., Léger, J.-M., Dolain, C., Guionneau, P., and Huc, I. (2003). Aromatic  $\delta$ -peptides: Design, synthesis and structural studies of helical, quinoline-derived oligoamide foldamers. **Tetrahedron** 59, 8365–8374.
- Johnson, L.M., and Gellman, S.H. (2013).  $\alpha$ -Helix mimicry with  $\alpha/\beta$ -peptides. **Meth. Enzymol.** 523, 407–429.
- Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. **Proc. Natl. Acad. Sci. U.S.A.** 93, 13–20.
- Kabsch, W. (2010). XDS. **Acta Crystallogr. D Biol. Crystallogr.** 66, 125–132.
- Kakkar, T., Sung, C., Gibiansky, L., Vu, T., Narayanan, A., Lin, S.-L., Vincent, M., Banfield, C., Colbert, A., Hoofring, S., et al. (2011). Population PK and IgE pharmacodynamic analysis of a fully human monoclonal antibody against IL4 receptor. **Pharm. Res.** 28, 2530–2542.
- Kantardjieff, K.A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. **Protein Sci.** 12, 1865–1871.
- Kendhale, A.M., Poniman, L., Dong, Z., Laxmi-Reddy, K., Kauffmann, B., Ferrand, Y., and Huc, I. (2011). Absolute control of helical handedness in quinoline oligoamides. **J. Org. Chem.** 76, 195–200.
- König, P., and Richmond, T.J. (1993). The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. **J. Mol. Biol.** 233, 139–154.
- Kraich, M., Klein, M., Patiño, E., Harrer, H., Nickel, J., Sebald, W., and Mueller, T.D. (2006). A modular interface of IL-4 allows for scalable affinity without affecting specificity for the IL-4 receptor. **BMC Biol.** 4, 13.
- Krishnamurthy, V.M., Kaufman, G.K., Urbach, A.R., Gitlin, I., Gudiksen, K.L., Weibel, D.B., and Whitesides, G.M. (2008). Carbonic anhydrase as a model for biophysical and physical-organic studies of proteins and protein-ligand binding. **Chem. Rev.** 108, 946–1051.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. **J. Mol. Biol.** 372, 774–797.

- Kritzer, J.A., Hodsdon, M.E., and Schepartz, A. (2005a). Solution structure of a beta-peptide ligand for hDM2. **J. Am. Chem. Soc.** *127*, 4118–4119.
- Kritzer, J.A., Stephens, O.M., Guarracino, D.A., Reznik, S.K., and Schepartz, A. (2005b). beta-Peptides as inhibitors of protein-protein interactions. **Bioorg. Med. Chem.** *13*, 11–16.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature** *227*, 680–685.
- LaPorte, S.L., Juo, Z.S., Vaclavikova, J., Colf, L.A., Qi, X., Heller, N.M., Keegan, A.D., and Garcia, K.C. (2008). Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system. **Cell** *132*, 259–272.
- Lee, E.F., Sadowsky, J.D., Smith, B.J., Czabotar, P.E., Peterson-Kaufman, K.J., Colman, P.M., Gellman, S.H., and Fairlie, W.D. (2009). High-resolution structural characterization of a helical alpha/beta-peptide foldamer bound to the anti-apoptotic protein Bcl-xL. **Angew. Chem. Int. Ed. Engl.** *48*, 4318–4322.
- Leslie, A.G.W. (2006). The integration of macromolecular diffraction data. **Acta Crystallogr. D Biol. Crystallogr.** *62*, 48–57.
- Letzelter, F., Wang, Y., and Sebald, W. (1998). The interleukin-4 site-2 epitope determining binding of the common receptor gamma chain. **Eur. J. Biochem.** *257*, 11–20.
- Li, C., Liu, M., Monbo, J., Zou, G., Li, C., Yuan, W., Zella, D., Lu, W.-Y., and Lu, W. (2008). Turning a scorpion toxin into an antitumor miniprotein. **J. Am. Chem. Soc.** *130*, 13546–13548.
- Liang, S., Dai, J., Hou, S., Su, L., Zhang, D., Guo, H., Hu, S., Wang, H., Rao, Z., Guo, Y., et al. (2013). Structural basis for treating tumor necrosis factor  $\alpha$  (TNF $\alpha$ )-associated diseases with the therapeutic antibody infliximab. **J. Biol. Chem.** *288*, 13799–13807.
- Liljas, A., Kannan, K.K., Bergstén, P.C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Järup, L., Lövgren, S., and Petef, M. (1972). Crystal structure of human carbonic anhydrase C. **Nature: New Biology** *235*, 131–137.
- Lind, R., Greenhow, D., Perry, S., Kimmerlin, T., and Seebach, D. (2004). Comparative metabolism of alpha- and beta-peptides in the insect *Heliothis virescens* and in plant cells of black Mexican sweet maize. **Chem. Biodivers.** *1*, 1391–1400.
- Luzina, I.G., Keegan, A.D., Heller, N.M., Rook, G.A.W., Shea-Donohue, T., and Atamas, S.P. (2012). Regulation of inflammation by interleukin-4: a review of “alternatives.” **J. Leukoc. Biol.** *92*, 753–764.
- Martin, L., Stricher, F., Missé, D., Sironi, F., Pugnière, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., et al. (2003). Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. **Nat. Biotechnol.** *21*, 71–76.
- Massover, W.H. (2007). Radiation damage to protein specimens from electron beam imaging and diffraction: a mini-review of anti-damage approaches, with special reference to synchrotron X-ray crystallography. **J Synchrotron Radiat** *14*, 116–127.



- Matthews, T., Salgo, M., Greenberg, M., Chung, J., DeMasi, R., and Bolognesi, D. (2004). Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. **Nat Rev Drug Discov** 3, 215–225.
- Maupin, C.M., McKenna, R., Silverman, D.N., and Voth, G.A. (2009). Elucidation of the proton transport mechanism in human carbonic anhydrase II. **J. Am. Chem. Soc.** 131, 7598–7608.
- McPherson, A. (1982). Preparation and analysis of protein crystals. **John Wiley and Sons, Inc.**
- Meazza, R., Azzarone, B., Orengo, A.M., and Ferrini, S. (2011). Role of common-gamma chain cytokines in NK cell development and function: perspectives for immunotherapy. **J. Biomed. Biotechnol.** 2011, 861920.
- Moreira, I.S., Fernandes, P.A., and Ramos, M.J. (2007). Hot spots--a review of the protein-protein interface determinant amino-acid residues. **Proteins** 68, 803–812.
- Moy, F.J., Diblasio, E., Wilhelm, J., and Powers, R. (2001). Solution structure of human IL-13 and implication for receptor binding. **J. Mol. Biol.** 310, 219–230.
- Mueller, T.D., Zhang, J.-L., Sebald, W., and Duschl, A. (2002). Structure, binding, and antagonists in the IL-4/IL-13 receptor system. **Biochim. Biophys. Acta** 1592, 237–250.
- Müller, M.M., Windsor, M.A., Pomerantz, W.C., Gellman, S.H., and Hilvert, D. (2009). A rationally designed aldolase foldamer. **Angew. Chem. Int. Ed. Engl.** 48, 922–925.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. **Acta Crystallogr. D Biol. Crystallogr.** 53, 240–255.
- Myszka, D.G. (1999). Improving biosensor analysis. **J. Mol. Recognit.** 12, 279–284.
- Nelms, K., Keegan, A.D., Zamorano, J., Ryan, J.J., and Paul, W.E. (1999). The IL-4 receptor: signaling mechanisms and biologic functions. **Annu. Rev. Immunol.** 17, 701–738.
- Nishimori, I., Minakuchi, T., Kohsaki, T., Onishi, S., Takeuchi, H., Vullo, D., Scozzafava, A., and Supuran, C.T. (2007). Carbonic anhydrase inhibitors: the beta-carbonic anhydrase from *Helicobacter pylori* is a new target for sulfonamide and sulfamate inhibitors. **Bioorg. Med. Chem. Lett.** 17, 3585–3594.
- Nooren, I.M.A., and Thornton, J.M. (2003). Diversity of protein-protein interactions. **EMBO J.** 22, 3486–3492.
- Oh, C.K., Geba, G.P., and Molfino, N. (2010). Investigational therapeutics targeting the IL-4/IL-13/STAT-6 pathway for the treatment of asthma. **Eur Respir Rev** 19, 46–54.
- Ohara, J., and Paul, W.E. (1988). Up-regulation of interleukin 4/B-cell stimulatory factor 1 receptor expression. **Proc. Natl. Acad. Sci. U.S.A.** 85, 8221–8225.
- Osborne, S.E., and Ellington, A.D. (1997). Nucleic Acid Selection and the Challenge of Combinatorial Chemistry. **Chem. Rev.** 97, 349–370.
- Palau, W., and Di Primo, C. (2013). Simulated single-cycle kinetics improves the design of surface

- p lasmon resonance assays.
- Talanta**
- 114*
- , 211–216.
- Papalia, G.A., Leavitt, S., Bynum, M.A., Katsamba, P.S., Wilton, R., Qiu, H., Steukers, M., Wang, S., Bindu, L., Phogat, S., et al. (2006). Comparative analysis of 10 small molecules binding to carbonic anhydrase II by different investigators using Biacore technology. **Anal. Biochem.** *359*, 94–105.
- Pocker, Y., and Sarkanen, S. (1978). Oxonase and esterase activities of erythrocyte carbonic anhydrase. **Biochemistry** *17*, 1110–1118.
- Renz, H., Bradley, K., Enssle, K., Loader, J.E., Larsen, G.L., and Gelfand, E.W. (1996). Prevention of the development of immediate hypersensitivity and airway hyperresponsiveness following in vivo treatment with soluble IL-4 receptor. **Int. Arch. Allergy Immunol.** *109*, 167–176.
- Reusch, P., Arnold, S., Heusser, C., Wagner, K., Weston, B., and Sebald, W. (1994). Neutralizing monoclonal antibodies define two different functional sites in human interleukin-4. **Eur. J. Biochem.** *222*, 491–499.
- Riddles, P.W., Blakeley, R.L., and Zerner, B. (1983). Reassessment of Ellman's reagent. **Meth. Enzymol.** *91*, 49–60.
- Robinson, J.A. (2009). Design of protein-protein interaction inhibitors based on protein epitope mimetics. **Chembiochem** *10*, 971–973.
- Russell, S.M., Keegan, A.D., Harada, N., Nakamura, Y., Noguchi, M., Leland, P., Friedmann, M.C., Miyajima, A., Puri, R.K., and Paul, W.E. (1993). Interleukin-2 receptor gamma chain: a functional component of the interleukin-4 receptor. **Science** *262*, 1880–1883.
- Sadowsky, J.D., Fairlie, W.D., Hadley, E.B., Lee, H.-S., Umezawa, N., Nikolovska-Coleska, Z., Wang, S., Huang, D.C.S., Tomita, Y., and Gellman, S.H. (2007). (alpha/beta+alpha)-peptide antagonists of BH3 domain/Bcl-x(L) recognition: toward general strategies for foldamer-based inhibition of protein-protein interactions. **J. Am. Chem. Soc.** *129*, 139–154.
- Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U., and Eisenberg, D. (2004). The Database of Interacting Proteins: 2004 update. **Nucleic Acids Res.** *32*, D449–451.
- Saraogi, I., and Hamilton, A.D. (2008). alpha-Helix mimetics as inhibitors of protein-protein interactions. **Biochem. Soc. Trans.** *36*, 1414–1417.
- Sebald, W., Nickel, J., Zhang, J.-L., and Mueller, T.D. (2010). Molecular basis of cytokine signalling--theme and variations. **FEBS J.** *277*, 106–118.
- Seder, R.A., Paul, W.E., Davis, M.M., and Fazekas de St Groth, B. (1992). The presence of interleukin 4 during in vitro priming determines the lymphokine-producing potential of CD4+ T cells from T cell receptor transgenic mice. **J. Exp. Med.** *176*, 1091–1098.
- Seebach, D., and Gardiner, J. (2008). Beta-peptidic peptidomimetics. **Acc. Chem. Res.** *41*, 1366–1375.
- Seebach, D., Beck, A.K., and Bierbaum, D.J. (2004). The world of beta- and gamma-peptides comprised of homologated proteinogenic amino acids and other components. **Chem. Biodivers.** *1*, 1111–1239.
- Sheridan, R.P., Deakyne, C.A., and Allen, L.C. (1980). Acetaldehyde hydrate and carbonic anhydrase:

- possible roles in the inhibition of brain aldehyde dehydrogenase. **Adv. Exp. Med. Biol.** *132*, 705–713.
- Smith, B.J., Lee, E.F., Checco, J.W., Evangelista, M., Gellman, S.H., and Fairlie, W.D. (2013). Structure-Guided Rational Design of  $\alpha/\beta$ -Peptide Foldamers with High Affinity for BCL-2 Family Prosurvival Proteins. **Chembiochem.** *14*, 1564–1572.
- Smith, L.J., Redfield, C., Boyd, J., Lawrence, G.M., Edwards, R.G., Smith, R.A., and Dobson, C.M. (1992). Human interleukin 4. The solution structure of a four-helix bundle protein. **J. Mol. Biol.** *224*, 899–904.
- Sokol, C.L., Barton, G.M., Farr, A.G., and Medzhitov, R. (2008). A mechanism for the initiation of allergen-induced T helper type 2 responses. **Nat. Immunol.** *9*, 310–318.
- Stockwin, L.H., and Holmes, S. (2003). Antibodies as therapeutic agents: vive la renaissance! **Expert Opin Biol Ther** *3*, 1133–1152.
- Stoltenburg, R., Reinemann, C., and Strehlitz, B. (2007). SELEX--a (r)evolutionary method to generate high-affinity nucleic acid ligands. **Biomol. Eng.** *24*, 381–403.
- Tina, K.G., Bhadra, R., and Srinivasan, N. (2007). PIC: Protein Interactions Calculator. **Nucleic Acids Res.** *35*, W473–476.
- Tony, H.P., Shen, B.J., Reusch, P., and Sebald, W. (1994). Design of human interleukin-4 antagonists inhibiting interleukin-4-dependent and interleukin-13-dependent responses in T-cells and B-cells with high efficiency. **Eur. J. Biochem.** *225*, 659–665.
- Tuncbag, N., Gursoy, A., and Keskin, O. (2009). Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. **Bioinformatics** *25*, 1513–1520.
- Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. **Acta Crystallogr. D Biol. Crystallogr.** *66*, 22–25.
- Vasilescu, A., Heath, S.C., Ivanova, R., Hendel, H., Do, H., Mazoyer, A., Khadivpour, E., Goutalier, F.X., Khalili, K., Rappaport, J., et al. (2003). Genomic analysis of Th1-Th2 cytokine genes in an AIDS cohort: identification of IL4 and IL10 haplotypes associated with the disease progression. **Genes Immun.** *4*, 441–449.
- Vasudev, P.G., Chatterjee, S., Shamala, N., and Balaram, P. (2009). Gabapentin: a stereochemically constrained gamma amino acid residue in hybrid peptide design. **Acc. Chem. Res.** *42*, 1628–1639.
- Vaz, D.C., Rodrigues, J.R., Sebald, W., Dobson, C.M., and Brito, R.M.M. (2006). Enthalpic and entropic contributions mediate the role of disulfide bonds on the conformational stability of interleukin-4. **Protein Sci.** *15*, 33–44.
- Verdine, G.L., and Walensky, L.D. (2007). The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. **Clin. Cancer Res.** *13*, 7264–7270.
- Veselovsky, A.V., Ivanov, Y.D., Ivanov, A.S., Archakov, A.I., Lewi, P., and Janssen, P. (2002). Protein-protein interactions: mechanisms and modification by drugs. **J. Mol. Recognit.** *15*, 405–416.

422.

- Weber, A., Casini, A., Heine, A., Kuhn, D., Supuran, C.T., Scozzafava, A., and Klebe, G. (2004). Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. **J. Med. Chem.** *47*, 550–557.
- Wenzel, S., Wilbraham, D., Fuller, R., Getz, E.B., and Longphre, M. (2007). Effect of an interleukin-4 variant on late phase asthmatic response to allergen challenge in asthmatic patients: results of two phase 2a studies. **Lancet** *370*, 1422–1431.
- Werder, M., Hauser, H., Abele, S., Seebach, D. (1999).  $\beta$ -Peptides as Inhibitors of Small Intestinal Cholesterol and Fat Absorption. **Helv. Chim. Acta** *82*, 1774–1783.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. **Acta Crystallogr. D Biol. Crystallogr.** *67*, 235–242.
- Wlodawer, A., Wlodawer, A., Pavlovsky, A., and Gustchina, A. (1992). Crystal structure of human recombinant interleukin-4 at 2.25 Å resolution. **FEBS Lett.** *309*, 59–64.
- Xenarios, I., and Eisenberg, D. (2001). Protein interaction databases. **Curr. Opin. Biotechnol.** *12*, 334–339.
- Yin, H., and Hamilton, A.D. (2005). Strategies for targeting protein-protein interactions with synthetic agents. **Angew. Chem. Int. Ed. Engl.** *44*, 4130–4163.
- Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. **Science** *292*, 883–896.
- Zhang, D.-W., Zhao, X., Hou, J.-L., and Li, Z.-T. (2012). Aromatic amide foldamers: structures, properties, and functions. **Chem. Rev.** *112*, 5271–5316.
- Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. **Cell** *98*, 811–824.
- Zhang, J.-L., Simeonowa, I., Wang, Y., and Sebald, W. (2002). The high-affinity interaction of human IL-4 and the receptor  $\alpha$  chain is constituted by two independent binding clusters. **J. Mol. Biol.** *315*, 399–407.



# Production scientifique

## *Communications orales en anglais*

- “Recognition of protein surfaces by aromatic amide foldamers.” *Journée Scientifique du CBMN*, **Décembre 2012**, Pessac - France
- “Recognition of protein surfaces by aromatic amide foldamers - Application to human Carbonic Anhydrase II.” *High Throughput Protein Production and Crystallization Practical Course*, **Mai 2013**, Harwell, Grande-Bretagne

## *Posters en anglais*

- “Aromatic Amide foldamer-based recognition of protein surface.” *3<sup>rd</sup> International School on Biological Crystallization*, **Mai 2011**, Grenade - Espagne
- “Synthesis and evaluation of foldamers as inhibitors of carbonic anhydrase II” *Symposium on foldamers*, **Avril 2013**, Paris - France
- “Recognition of protein surfaces by aromatic amide foldamers - Application to human Carbonic Anhydrase II.” *High Throughput Protein Production and Crystallization Practical Course*, **Mai 2013**, Harwell, Grande-Bretagne
- “Reconnaissance de surfaces protéiques par des foldamères aromatiques.” *Congrès de l'Association Française de Cristallographie*, **Juin 2013**, Pessac - France.
- “Reconnaissance de surfaces protéiques par des foldamères aromatiques.” *Journée Scientifique du CBMN*, **Novembre 2013**, Pessac - France.

## *Communications écrites en anglais*

- Structure at 2.1 Å resolution of a helical aromatic oligoamide foldamer-protein complex. Manuscrit soumis dans *Angewandte Chemie* le **20 Octobre 2013** accepté le **4 Novembre 2013**.



# Structure at 2.1 Å resolution of a helical aromatic oligoamide foldamer-protein complex

Jérémie Buratto, Cinzia Colombo, Marine Stupfel, Simon J. Dawson, Christel Dolain, Béatrice Langlois d'Estaintot, Lucile Fischer, Thierry Granier, Michel Laguerre, Bernard Gallois\*, Ivan Huc\*

In memory of Dr. Frédéric Denonne

Most proteins operate as complexes or within networks of interactions with other proteins.<sup>[1]</sup> The recognition of protein surfaces and the inhibition of protein-protein interactions (PPIs) thus offer innumerable opportunities for the development of pharmacological tools or therapeutic agents. Yet they have remained extremely challenging objectives. Indeed, protein surfaces only rarely possess grooves or well-defined active sites that could be targeted by small molecules.<sup>[2]</sup> In recent years, PPI inhibitors have thus been developed using large molecules such as other proteins,<sup>[3]</sup> and in particular antibodies,<sup>[4]</sup> or aptamers.<sup>[5]</sup> Synthetic foldamers,<sup>[6]</sup> due to their medium size (typically in the 0.5-5 kDa range) and well-defined structure in solution, appear as potent candidates to serve as scaffolds bearing proteinogenic side chains that would recognize protein surfaces.<sup>[7]</sup> In addition, foldamers generally show high resistance towards enzyme degradation and represent attractive alternatives to peptidic and oligonucleotidic backbones. But the question remains: how to arrange individual proteinogenic side chains at the surface of a foldamer to target a given protein surface? Efforts towards this goal have largely concerned the development of  $\alpha$ -helix mimetics,<sup>[8]</sup> combining structure-based design and the screening of small foldamer libraries.<sup>[9]</sup> But the targeting of large surfaces not complementary to a discrete protein secondary motif remains elusive. A major hurdle lies in the fact that detailed

structural information about interactions between foldamers, or other medium size molecules, and proteins is very scarce.<sup>[8], [80], [10]</sup> In the absence of good lead compounds to form complexes that could be subjected to structural analysis and also in the absence of structural information to help designing good ligands, the discovery of foldamers to bind protein surfaces meets a sort of deadlock.

As a way to solve this problem, we reasoned that it may be possible to obtain structural information about interactions at a foldamer-protein interface even in the absence of strong binding, provided some sort of attachment would link the two. For this purpose, helical aromatic oligoamide foldamers<sup>[11]</sup> based on 8-amino-2-quinolinecarboxylic acid (**Q**, see Chart 1b) appeared to be suitable. They feature highly stable conformations<sup>[12]</sup> in particular in protic media.<sup>[13]</sup> Monomers can be equipped with proteinogenic side chains and oligomers can be prepared using solid phase synthesis (SPS).<sup>[14]</sup> Though they have not yet been shown to interact with proteins, tight binding to G-quadruplex DNA has been demonstrated.<sup>[15]</sup> In addition, helical aromatic oligoamides possess a remarkable ability to form single crystals even when their size exceeds 5 kDa,<sup>[16]</sup> allowing the hope of crystallizing protein-foldamer conjugates. We thus endeavored to prepare foldamers tethered to a protein ligand in order to crystallize a protein-ligand-foldamer complex and find a structural basis from which to tailor foldamer-protein interactions. As a relevant background to this study, previous examples had been described of the exploration of protein binding properties via the use of recognition groups, including large peptides, attached to a protein ligand.<sup>[17]</sup> In the following, we report the design and 2.1 Å resolution structure elucidation of a helical aromatic oligoamide foldamer-protein complex mediated by a protein ligand.

Human carbonic anhydrase II (HCA) was selected as a model target because it is commercial, relatively easy to crystallize and because structurally simple benzene-sulfonamide nanomolar inhibitors can be readily prepared.<sup>[17], [18]</sup> Benzyl 4-aminosulfonylbenzamide **1** (Chart 1) has a  $K_D$  value of *ca* 2 nM<sup>[17b], [19]</sup> and has been co-crystallized in complex with HCA site, with its sulfonyl group coordinated to the Zn(II) ion of the enzyme.<sup>[19]</sup> Examination of this structure hinted at the preference for a functionalization of the benzyl group in the *meta* position to extend this inhibitor out of the binding pocket and covalently attach a foldamer. We prepared inhibitor **2** (see Supporting Information) and confirmed its suitability by a crystal structure at 1 Å resolution of its complex with HCA, which crystallized in the classical  $P2_1$  lattice and showed the phenol function pointing straight out of the binding site (protein database entry: 4MTY). We thus prepared inhibitor **3**, possessing an aliphatic amine at the end of an *n*-butyloxy linker, which was inserted to allow the foldamer some freedom to position itself at the protein surface. A method was then developed to attach the aliphatic amine of **3** to the N-terminal aromatic amine of a quinoline oligoamide chain on the solid phase. The terminal 8-aminoquinoline

[\*] J. Buratto,<sup>†</sup> Dr. C. Colombo,<sup>†</sup> M. Stupfel, Dr. S. J. Dawson, Dr. C. Dolain, Dr. B. Langlois d'Estaintot, Dr. L. Fischer, Dr. T. Granier, Dr. M. Laguerre, Dr. B. Gallois, Dr. I. Huc  
Université de Bordeaux, CBMN, UMR 5248  
Institut Européen de Chimie Biologie  
2 rue Robert Escarpit 33607 Pessac, France  
and:  
CNRS, CBMN, UMR5248  
and:  
Institut Polytechnique de Bordeaux, UMR5248  
E-mail: b.gallois@cbmn.u-bordeaux.fr, i.huc@iecb.u-bordeaux.fr

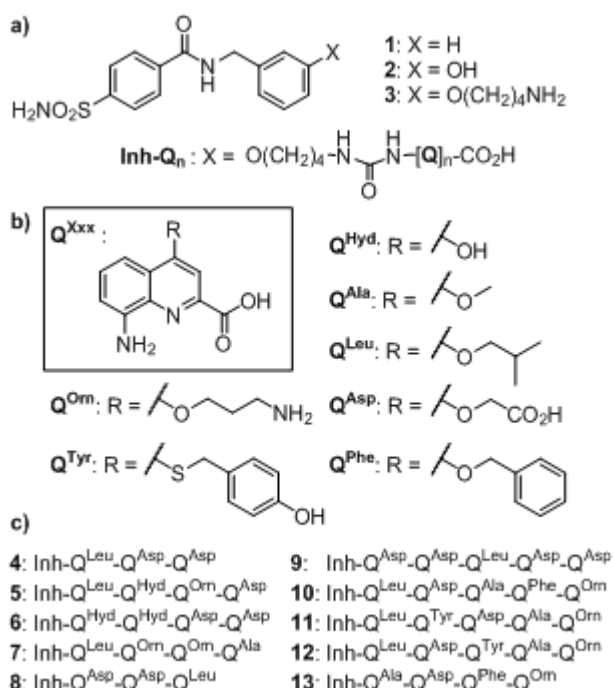
[\*\*] This work was supported by European Grant FP7-IAPP-2008-230662-Foldappi (Post-doctoral fellowships to C.C. and S.D.) and by the French Ministry of Research (Pre-doctoral fellowships to J. B. and M. S.). SPR measurements were carried out with the help of C. Di Primo on the technical platform of IECB (UMS3033/US001). We thank Dr. T. Deschrijver for providing precursors of Q<sup>Tyr</sup> and Q<sup>Hyd</sup> and Dr. B. Kauffmann for synchrotron diffraction measurements of the protein structures. We thank the ESRF for provision of synchrotron radiation facilities and the staff of the ID 23-2 beamline for their kind assistance during data collections.

<sup>†</sup> These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.







**Chart 1.** a) Inhibitor constructs; b)  $\text{Q}^{\text{xxx}}$  monomers; c) Inh-Q<sub>n</sub> foldamers synthesized by SPS.

was activated on the Wang resin using triphosgene and amine **3** was then added to form a urea. In the following, we designate such inhibitor-urea-oligoamide constructs as Inh-Q<sub>n</sub> (Chart 1a).

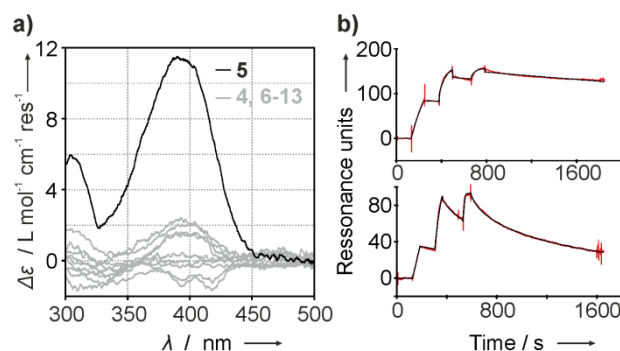
Proteinogenic side chains were attached by ether or thioether functions to the quinoline position 4 which diverge from the helices. Monomers are designated  $\text{Q}^{\text{xxx}}$  using the Xxx three-letter code of analogous  $\alpha$ -amino acids when available. All **Fmoc-Q<sup>xxx</sup>** monomers were prepared on a multigram scale in 4-5 steps and overall yields ranging from 20 to 60%, from a common 8-nitro-4-(1*H*)-quinolinone precursor. For the purpose of this study, we used the previously described **Fmoc-Q<sup>Leu</sup>**, **Fmoc-Q<sup>Asp</sup>**, and **Fmoc-Q<sup>Orn</sup>** acid precursors, the latter two having their side chain protected as the acid labile tBu ester and Boc amine, respectively.<sup>[14]</sup> We also introduced four new residues (see supporting information). To prepare **Fmoc-Q<sup>Phe</sup>**, the classical H<sub>2</sub>/Pd-C reduction of the 8-nitro-quinoline precursor to produce the corresponding amine was proscribed by the presence of the benzyl ether function. CuSO<sub>4</sub>/NaBH<sub>4</sub> proved to be an efficient alternative. Monomer  $\text{Q}^{\text{Hyd}}$  possesses a simple 4-hydroxy residue that was protected by a TFA labile *para*-methoxy-benzyl (PMB) ether during SPS. The PMB group resisted H<sub>2</sub>/Pd-C reduction of the 8-nitro-quinoline in the presence of pyridine as described in the literature.<sup>[20]</sup> In the case of **Fmoc-Q<sup>Tyr</sup>**, an acid stable thioether was used to attach the side chain to the quinoline, formed by aromatic nucleophilic substitution of a 4-bromoquinoline precursor by *para*-(*tert*-butoxy)-benzylmercaptan. The hydrogenation of the 8-nitro group was possible in this case too, but somewhat sluggish using the H<sub>2</sub>/Pd-C method possibly due to poisoning of the catalyst by sulfur derivatives. Use of NH<sub>4</sub>HCO<sub>2</sub>/Pd-C resolved this problem. **Fmoc-Q<sup>Ala</sup>** caused no particular difficulty and indeed, the final compound could be afforded in excellent purity via a simple precipitation.

In order to screen different foldamers for their interactions with the surface of HCA, ten sequences were initially prepared by SPS,<sup>[14]</sup> all having an appended HCA inhibitor (Chart 1c). Crude oligomers are typically obtained in 60-90 % purity. Side chain

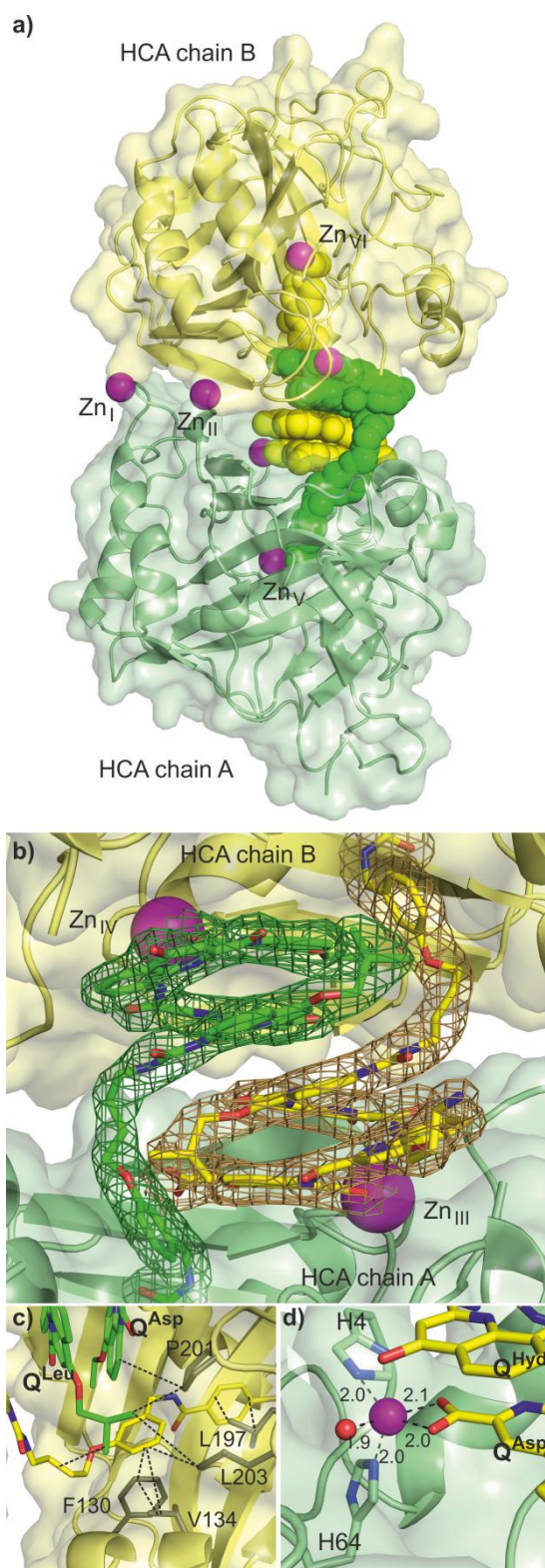
deprotection and cleavage from the resin was carried out in 95/2.5/2.5 TFA:*i*Pr<sub>3</sub>SiH:H<sub>2</sub>O vol/vol/vol. Sequences were purified by reversed phase HPLC and obtained in 30-80 % yield from initial Wang resin loadings. The choice of side chains of **4-13** was intended to be diverse, with most sequences carrying a mix of charged and hydrophobic residues, yet with a prevalence of  $\text{Q}^{\text{Asp}}$  which was favored due to the abundance of positively charged residues in the vicinity of the HCA active site.

Importantly, sequences **4-13** do not carry any stereogenic center and thus exist as a mixture of right- (*P*) and left- (*M*) handed enantiomeric helical conformers. **Q<sub>n</sub>** sequences were kept short (3 to 5 residues) to allow helix handedness inversion to occur in solution. Indeed, helix stability rises so quickly upon increasing oligomer length<sup>[12b]</sup> that longer sequences do not undergo any equilibrium in protic media.<sup>[13]</sup> We then used circular dichroism (CD) as a preliminary screen. All compounds **4-13** were expected to have a high affinity for HCA imparted by their inhibitor moiety. However, the desired interactions between the foldamer helix and the protein, if they existed, would be diastereoselective and thus induce a preferred handedness in the helical oligomer backbone.<sup>[15a,21]</sup>

The CD spectra of **4-13** were recorded in the presence of HCA and monitored over time during 2-5 days to allow helix handedness inversion to occur. The time necessary for the CD signal to stabilize was dependent on the length of the foldamers. Two days were necessary for sequences containing 3 or 4 monomers (**4-8**; **13**). Five days were necessary for sequence bearing 5 monomers (**9-12**). HCA does not absorb between 300-500 nm, and thus any CD band could be assigned to the quinoline chromophores. As shown in Fig. 1a, all CD spectra but one showed very weak bands indicating only marginal effects of the protein surface on foldamer handedness. In contrast, an intense CD band emerged at 390 nm in the spectrum of **5**. Comparison of the  $\Delta\epsilon_{390}$  value (40 Lmol<sup>-1</sup>cm<sup>-1</sup> for 4 residues) with those of related oligomers<sup>[22]</sup> suggested that an almost complete handedness bias of *5* towards a *P* helix had taken place upon confinement to the protein surface. The absence of induced CD for the other compounds does not allow us to rule out strong foldamer-protein interactions, but reveals no selectivity between *P* and *M* helices. Compounds **4** and **6-13** were thus not further investigated.



**Figure 1.** a) CD spectra of **4-13** (34.5  $\mu\text{M}$  in 50 mM aqueous phosphate, pH 7.4) in the presence of HCA (34.5  $\mu\text{M}$ ) at 20°C, at equilibrium (2 or 5 days of incubation); b) SPR sensorgrams (in red) in 2:98 DMSO:aqueous PBS pH 7.4 (vol/vol) at 25 °C. Increasingly concentrated solutions of **5** (250-500-1000 nM, top) and of **3** (62.5-125-250 nM, bottom) were flowed over the HCA surface sequentially at 25  $\mu\text{L}/\text{min}$  for 60 s. Curve fitting (in black) assuming a Langmuir 1:1 model yielded  $k_{\text{on}} = 2.8 \cdot 10^4 \text{ M}^{-1} \text{ s}^{-1}$  and  $k_{\text{off}} = 1.5 \cdot 10^{-4} \text{ s}^{-1}$  for foldamer **5**;  $k_{\text{on}} = 1.5 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ ,  $k_{\text{off}} = 7.7 \cdot 10^{-3} \text{ s}^{-1}$  for compound **3**.



**Figure 2.** Crystal structure of the HCA:5 complex. a) Asymmetric unit showing two HCA molecules in pale yellow (chain A) and green (chain B), two **Inh-Q<sup>Leu</sup>Q<sup>Hyd</sup>Q<sup>Orn</sup>Q<sup>Asp</sup>** molecules (**5**) in CPK representation and Zn<sup>2+</sup> ions as purple spheres. Zn<sub>I</sub> and Zn<sub>II</sub> are involved in interactions between HCA chains A and B. Zn<sub>III</sub> and Zn<sub>IV</sub> take part in protein-foldamer interactions. Zn<sub>V</sub> and Zn<sub>VI</sub> belong to catalytic centers; b) Foldamer backbones shown in sticks and contoured by 2mFo-DFc density maps at 1 $\sigma$  level; c) contacts between the hydrophobic side chain of **Q<sup>Leu</sup>** (green), the linker of the **Inh** of the other foldamer (yellow) and hydrophobic HCA residues (grey); d) Zn<sub>III</sub> (purple sphere) and its ligands. A neighbor **Q<sup>Hyd</sup>** hydroxyl group is found in a distal position. Distances are shown in Å.

The binding of **5** to HCA was then characterized by surface plasmon resonance (SPR) using established protocols.<sup>[23]</sup> Kinetic titration assays<sup>[24]</sup> were first performed with model compound **3** on HCA immobilized on the chip, allowing a  $K_D$  value of *ca.* 5 nM to be determined (Fig. 1). Foldamer **5** showed an almost identical  $K_D$  value but its  $k_{on}$  and  $k_{off}$  rates were both 50 times smaller than for **3**. This slow binding and dissociation confirms that interactions take place between the foldamer and HCA, independent from the **Inh** part of **5**.

Cocrystals of the HCA-**5** complex were obtained by the hanging drop technique and allowed its structure to be solved at 2.1 Å resolution (protein database entry: 4LP6, see Supporting Information). The complex crystallizes in space group  $P2_1$  but the packing is unprecedented among over 400 reported HCA structures. The unit cell volume is twice that of the search model and of the HCA-**2** complex, so that two independent HCA-**5** subunits are present in the asymmetric unit (Fig. 2a). The final model shows a continuous (2mFo-DFc) electron density map, contoured at 1.0  $\sigma$ , all along both A and B HCA chains (from residues 4 to 260) and the two well-defined **Inh-Q<sub>4</sub>** molecules (Fig. 2b). As with other HCA structures, no density is observed for the three first N-terminal residues of each chain.

The structure reveals the formation of a novel pseudo  $C_2$ -symmetrical HCA dimer (Fig. 2a) involving new protein-protein contacts, foldamer-protein interactions, and foldamer-foldamer interactions. Within a dimer, protein-protein interactions are mediated by two water molecules hydrogen bonded to Ser172 residues of chains A and B and by two Zn<sup>2+</sup> bridges (Fig. S2). Coordination spheres involve Lys171 and Glu233 of one chain and Asp174 of the other chain. For one of these bridges, an additional ligand, Glu26 from symmetry related chain B, completes the coordination sphere. Within the crystal lattice, contacts between dimers are in limited number, which may explain a high value (51.5 Å<sup>2</sup>) for the average temperature factor.

As expected, the inhibitor moieties are bound to the HCA active sites in a conventional manner, with the sulfonamide group coordinated to the catalytic zinc, the benzene-sulfonamide ring stacked over Leu197 and the benzyl ring in an edge-to-face arrangement with Phe130 (Fig. 2c). The foldamers protrude from the protein surface and adopt right-handed conformations, in agreement with CD spectra. The foldamer bound to chain A is more disordered (higher temperature factor values,  $Q^{Orn}$  and  $Q^{Asp}$  side chains not observed in density map) than that bound to chain B. The active sites of HCA chains A and B face each other allowing a direct contact between the foldamers that contribute to the dimerization. Remarkably, the inhibitor-foldamer linkers are in an extended conformation. As a result, each foldamer extends beyond the HCA chain to which its inhibitor moiety is bound and establishes interactions with the other HCA chain. In addition, the two foldamers stack to clip into one another. Stacking interactions occur through the first quinoline units in each sequence (**Q<sup>Leu</sup>**) which are found to be co-planar (average distance between rings: 3.46 Å) and rotated by almost 90°. The helix axes of the two foldamers deviate by only 7°.

At its C-terminal part, each helix interacts with the opposite HCA chain through a Zn<sup>2+</sup> cation coordinated in a square-based pyramidal geometry (Fig. 2d). The involved ligands are the main chain C-terminal carboxylate moiety of the last quinoline, the His4 and His64 imidazole rings and a water molecule. Cooperative metal binding by His4 and an HCA ligand had been hypothesized by others in earlier studies and is validated here.<sup>[10]</sup> The hydroxy group of the **Q<sup>Hyd</sup>** is found at a distal position in the Zn<sup>2+</sup> coordination



sphere (mean- $d_{\text{Zn-O}} = 3.8 \text{ \AA}$ ). The role of  $\text{Zn}^{2+}$  in both protein-protein and foldamer-protein contacts seems to be essential to crystal packing. Indeed, no crystal growth was observed in the absence of zinc acetate.

When observed (foldamer bound to chain B), the polar  $\text{Q}^{\text{Asp}}$  and  $\text{Q}^{\text{Orn}}$  side chains are found to be hydrogen bonded to the opposite protein chain *via* bridging water molecules. The  $\text{Q}^{\text{Leu}}$  side chain is the only one that is clearly observed for both foldamers and involved in direct interaction with the protein, as it faces a hydrophobic environment involving Pro201, Leu203, the linker and the phenyl ring of the inhibitor (Fig. 2c).

These multiple foldamer-protein and protein-protein interactions do not alter the overall HCA structure. Indeed, the A and B chains are similar to each other (average r.m.s deviation on 257 Ca positions: 0.23 Å; largest r.m.s deviation: 0.93 Å), and similar to the native protein (average and largest Ca r.m.s deviations when superimposing HCA-5 to the model are 0.31 Å and 0.95 Å for chain A, 0.33 Å and 1.14 Å for chain B, respectively). The reciprocal is also true, the foldamer main chain helix is not altered by the multiple interactions its side chain, inhibitor moiety and terminal carboxylate engage with the protein.

Altogether, the above results validate the strategy proposed in this study. Anchoring a foldamer to a modified inhibitor does not drastically change the  $K_D$  value which remains in the nanomolar range. Accurate structural information quickly revealed intricate interactions between foldamers, proteins and  $\text{Zn}^{2+}$  ions and an unexpected and unprecedented dimerization. Multiple suggestions for improvements and controls can be proposed from the structure of HCA-5. Can the  $\text{Q}^{\text{Orn}}$  and  $\text{Q}^{\text{Asp}}$  residues of **5** be modified to enhance foldamer-protein interactions? Can the  $\text{Q}^{\text{Hyd}}$  residue be modified to complete the coordination sphere of the foldamer-protein  $\text{Zn}^{2+}$  bridge? The crystals grew at a concentration of HCA-5 complex of 150  $\mu\text{M}$ , which hints at a reasonable stability of the dimeric structure. Yet, this stability and its dependence on  $\text{Zn}^{2+}$  concentration should be assessed in solution. Investigations along these lines are in progress and will be reported in due course.

Received: ((will be filled in by the editorial staff))

Published online on ((will be filled in by the editorial staff))

**Keywords:** Foldamers • protein surface recognition • human carbonic anhydrase • crystallography • solid phase synthesis

- [1] T. Berggard, S. Linse, P. James, *Proteomics* **2007**, 7, 2833; b) W. E. Stites, *Chem. Rev.* **1997**, 97, 1233.
- [2] J. A. Wells, C. L. McLendon, *Nature* **2007**, 450, 1001; T. Berg, *Angew. Chem.* **2003**, 115, 2566; *Angew. Chem. Int. Ed.* **2003**, 42, 2462.
- [3] For representative examples, see: a) M. G. Kolonin, R. L. Finley, Jr, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 14266; b) R. Woodman, J. T.-H. Yeh, S. Laurenson, P. K. Ferrigno, *J. Mol. Biol.* **2005**, 352, 1118; c) S. L. LaPorte, C. M. Forsythe, B. C. Cunningham, L. J. Miercke, D. Akhavan, R. M. Stroud, *Proc. Natl. Acad. Sci. USA*, **2005**, 102, 1889; d) T. Robson, I. F. James, *Drug Discover Today* **2012**, 17, 544; e) Y. Ji, S. Majumder, M. Millard, R. Borra, T. Bi, A. Y. Elnager, N. Neamati, A. Shekhtman, J. A. Camarero, *J. Am. Chem. Soc.* **2013**, 135, 11623.
- [4] For reviews of this subject, see: a) L. Sanz, B. Blanco, L. Álvarez-Vallina, *Trends Immunol.* **2004**, 25, 1471; b) P. Holliger, P. J. Hudson, *Nat. Biotechnol.* **2005**, 23, 1126; c) E. B. DeSouza, S. T. Cload, P. S. Pendergrast, D. W. Y. Sah, *Neuropsychopharmacology* **2009**, 34, 142; d) Á. M. Cuesta, N. Sainz-Pastor, J. Bonet, B. Oliva, L. Álvarez-Vallina, *Trends Biotechnol.* **2010**, 28, 355.
- [5] For representative examples, see: a) R. Conrad, L. M. Keranen, A. D. Ellington, A. C. Newton, *J. Biol. Chem.* **1994**, 269, 32051; b) R.-H. Huang, D. H. Fremont, J. L. Diener, R. G. Schaub, J. E. Sadler, *Structure*, **2009**, 17, 1476; c) Y. S. Choi, J. Hur, H. K. Lee, S. Jeong, *FEBS Lett.* **2009**, 583, 1415; d) J. B. Madsen, D. M. Dupont, T. B. Andersen, A. F. Nielsen, L. Sang, D. M. Brix, J. K. Jensen, T. Broos, M. L. V. Hendrickx, A. Christensen, J. Kjems, P. A. Andreasen, *Biochemistry* **2010**, 49, 4103; e) V. M. Tesmer, S. Lennarz, G. Mayer, J. J. G. Tesmer, *Structure*, **2012**, 20, 1300; f) T. Mashima, F. Nishikawa, Y. O. Kamatari, H. Fujiwara, M. Saimura, T. Nagata, T. Kodaki, S. Nishikawa, K. Kuwata, M. Katahira, *Nucleic Acids Res.* **2013**, 41, 1355.
- [6] G. Guichard, I. Huc, *Chem. Commun.* **2011**, 47, 5933.
- [7] a) H. Yin, A. D. Hamilton, *Angew. Chem.* **2005**, 117, 4200; *Angew. Chem. Int. Ed.* **2005**, 44, 4130; b) V. Azzarito, K. Long, N. S. Murphy, A. J. Wilson, *Nat. Chem.* **2013**, 161; B. Baptiste, F. Godde, I. Huc, *ChemBioChem* **2009**, 10, 1765.
- [8] For representative examples, see: a) H. S. Park, Q. Lin, A. D. Hamilton, *J. Am. Chem. Soc.* **1999**, 121, 8; b) O. Kutzki, H. S. Park, J. T. Ernst, B. P. Ormer, H. Yin, A. D. Hamilton, *J. Am. Chem. Soc.* **2002**, 124, 11838; c) J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem.* **2003**, 115, 553; *Angew. Chem. Int. Ed.* **2003**, 42, 535; d) J. A. Kritzer, J. A. Lear, M. E. Hodsdon, A. Schepartz, *J. Am. Chem. Soc.* **2004**, 126, 9468; e) J. D. Sadowsky, M. A. Schmitt, H.-S. Lee, N. Umezawa, S. Wang, Y. Tomita, S. H. Gellman, *J. Am. Chem. Soc.* **2005**, 127, 11966; f) H. Yin, G.-i. Lee, K. A. Sedey, J. M. Rodriguez, H.-G. Wang, S. M. Sebt, A. D. Hamilton, *J. Am. Chem. Soc.* **2005**, 127, 5463; g) R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, D. Obrecht, P. R. E. Mittl, M. G. Gruetter, J. A. Robinson, *ChemBioChem* **2006**, 7, 515; h) J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards, A. J. Wilson, *Chem. Commun.* **2009**, 5091; i) J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards, A. J. Wilson, *Chem. Commun.* **2009**, 5091; j) J. Michel, E. A. Harker, J. Tirado-Rives, W. L. Jorgensen, A. Schepartz, *J. Am. Chem. Soc.* **2009**, 131, 6356; k) J. A. Robinson, *ChemBioChem* **2009**, 10, 971; l) E. F. Lee, J. D. Sadowsky, B. J. Smith, P. E. Czabotar, K. J. Peterson-Kaufman, P. M. Colman, S. H. Gellman, W. D. Fairlie, *Angew. Chem.* **2009**, 121, 4382; *Angew. Chem. Int. Ed.* **2009**, 48, 4318; m) C. G. Cummings, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **2010**, 14, 341; n) I. Saraogi, J. A. Hebda, J. Becerril, L. A. Estroff, A. D. Miranker, A. D. Hamilton, *Angew. Chem.* **2010**, 122, 748; *Angew. Chem. Int. Ed.* **2010**, 49, 736; o) M. D. Boersma, H. S. Haase, K. J. Peterson-Kaufman, E. F. Lee, O. B. Clarke, P. M. Colman, B. J. Smith, S. W. Horne, W. D. Fairlie, S. H. Gellman, *J. Am. Chem. Soc.* **2012**, 134, 315.
- [9] a) W. S. Horne, M. D. Boersma, M. A. Windsor, S. H. Gellman, *Angew. Chem.* **2008**, 120, 2895; *Angew. Chem. Int. Ed.* **2008**, 47, 2853; b) A. Shaginian, L. R. Whitby, S. Hong, I. Hwang, B. Farooqi, M. Searcey, J. Chen, P. K. Vogt, D. L. Boger, *J. Am. Chem. Soc.* **2009**, 131, 5564.
- [10] a) J. T. Nguyen, C. W. Turck, F. E. Cohen, R. N. Zuckermann, W. A. Lim, *Science* **1998**, 282, 2088; b) R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem.* **2012**, 4, 527.
- [11] a) I. Huc, *Eur. J. Org. Chem.* **2004**, 17; b) D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* **2012**, 112, 5271.
- [12] (a) H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, 125, 3448; (b) N. Delsuc, T. Kawanami, J. Lefevre, A. Shundo, H. Ihara, M. Takafuji, I. Huc, *ChemPhysChem* **2008**, 9, 1882.
- [13] T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara, I. Huc, *Chem. Commun.* **2012**, 48, 6337.
- [14] B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *J. Org. Chem.* **2010**, 75, 7175.
- [15] a) P. S. Shirude, E. R. Gillies, S. Ladame, F. Godde, K. Shin-ya, I. Huc, S. Balasubramanian, *J. Am. Chem. Soc.* **2007**, 129, 11890; b) L. Delaurière, Z. Dong, K. Laxmi-Reddy, F. Godde, J.-J. Toulmé, I. Huc, *Angew. Chem.* **2012**, 124, 488; *Angew. Chem. Int. Ed.* **2012**, 51, 473.
- [16] a) N. Delsuc, J.-M. Léger, S. Massip, I. Huc, *Angew. Chem.* **2007**, 119, 218; *Angew. Chem. Int. Ed.* **2007**, 46, 214; b) N. Delsuc, S. Massip, J.-M. Léger, B. Kauffmann, I. Huc, *J. Am. Chem. Soc.* **2011**, 133, 3165.
- [17] a) A. Jain, S. G. Huang, G. M. Whitesides, *J. Am. Chem. Soc.* **1994**, 116, 5057; b) A. Jain, G. M. Whitesides, R. S. Alexander, D. W. Christianson, *J. Med. Chem.* **1994**, 37, 2100; c) T. Andersson, M.

- Lundquist, G. T. Dolphin, K. Enander, B.-H. Jonsson, J. W. Nilsson, L. Baltzer, *Chem. & Biol.* **2005**, *12*, 1245; d) L. T. Tegler, K. Fromell, B.-H. Jonsson, J. Viljanen, C. Winander, J. Carlsson, L. Baltzer, *ChemBioChem* **2011**, *12*, 559; e) K. Enander, G. T. Dolphin, L. Baltzer, *J. Am. Chem. Soc.* **2004**, *126*, 446; f) A. L. Banerjee, M. Swanson, B. C. Roy, X. Jia, M. K. Haldar, S. Mallik, D. K. Srivastava, *J. Am. Chem. Soc.* **2004**, *126*, 10875.
- [18] V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudiksen, D. B. Weibel, G. M. Whitesides, *Chem. Rev.* **2008**, *108*, 946.
- [19] C. Kim, J. S. Chang, J. B. Doyon, T. T. Baird, C. A. Fierke, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125.
- [20] H. Sajiki, H. Kuno, K. Hirota, *Tetrahedron Lett.* **1997**, *38*, 399
- [21] Y. Ferrand, A. M. Kendhale, B. Kauffmann, A. Grélaud, C. Marie, V. Blot, M. Pipelier, D. Dubreuil, I. Huc, *J. Am. Chem. Soc.* **2010**, *132*, 7858.
- [22] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195.
- [23] a) D.G. Myszka, *Analytical Biochemistry* **2004**, *329*, 316. b) G. A. Papalia et al., *Analytical Biochemistry* **2006**, *359*, 94.
- [24] a) R. Karlsson, P. S. Katsamba, H. Nordin, E. Pol, D. G. Myszka, *Analytical Biochemistry* **2006**, *349*, 136. b) C. Di Primo, *J. Mol. Recognit.* **2008**, *21*, 37. c) W. Palau, C. Masante, M. Ventura, C. Di Primo, *RNA* **2013**, *19*



